

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
31 July 2003 (31.07.2003)

PCT

(10) International Publication Number
WO 03/062400 A2

(51) International Patent Classification⁷: **C12N**

Knolls Parkway, Ijamsville, MD 21754 (US). **STEVEN-SON, Susan, C.** [US/US]; 10974 Horseshoe Drive, Frederick, MD 21701 (US).

(21) International Application Number: PCT/US03/02295

(22) International Filing Date: 24 January 2003 (24.01.2003)

(74) Agents: **SEIDMAN, Stephanie, L.** et al.; Heller Ehrman White & McAuliffe LLP, 7th Floor, 4350 La Jolla Village Drive, San Diego, CA 92122-1246 (US).

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

60/350,388 24 January 2002 (24.01.2002) US
60/391,967 26 June 2002 (26.06.2002) US

(63) Related by continuation (CON) or continuation-in-part (CIP) to earlier applications:

US 60/391,967 (CIP)
Filed on 26 June 2002 (26.06.2002)
US 60/350,388 (CIP)
Filed on 24 January 2002 (24.01.2002)

(71) Applicants (for all designated States except US): **THE SCRIPPS RESEARCH INSTITUTE** [US/US]; 10550 North Torrey Pines Road, TPC-8, La Jolla, CA 92037 (US). **NOVARTIS AG** [CH/CH]; Lichtstrasse 35, CH-4056 Basel (CH).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **KALEKO, Michael** [US/US]; 8 Heartstone Court, Rockville, MD 20854 (US). **NEMEROW, Glen, R.** [US/US]; 462 Cerro Street, Encinitas, CA 92024 (US). **SMITH, Theodore** [US/US]; 3346

(81) Designated States (*national*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States (*regional*): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, SE, SI, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— without international search report and to be republished upon receipt of that report

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: FIBER SHAFT MODIFICATIONS FOR EFFICIENT TARGETING

(57) Abstract: Provided are adenoviral vectors and the production of such vectors. In particular, fiber shaft modifications for efficient targeting of adenoviral vectors are provided. The fiber shaft modifications can be combined with other modifications, such as fiber knob and/or penton modifications, to produce fully ablated (detargeted) adenoviral vectors. A scale-up method for the propagation of detargeted adenoviral vectors is also provided.

WO 03/062400 A2

-1-

FIBER SHAFT MODIFICATIONS FOR EFFICIENT TARGETING RELATED APPLICATIONS

Benefit of priority is claimed to U.S. provisional application Serial No. 60/350,388, filed 24 January 2002, entitled "FIBER SHAFT MODIFICATIONS
5 FOR EFFICIENT TARGETING," to Stevenson, Susan C., Kaleko, Michael, Smith, Theodore and Nemerow, Glen R., and to U.S. provisional application Serial No. 60/391,967, filed 26 June 2002, entitled "FIBER SHAFT MODIFICATIONS FOR
10 EFFICIENT TARGETING," to Stevenson, Susan C., Kaleko, Michael, Smith, Theodore and Nemerow, Glen R. This application is also related to International PCT application No. (attorney docket number 22908-1236), filed the same day
herewith, entitled "FIBER SHAFT MODIFICATIONS FOR EFFICIENT
TARGETING," to Stevenson, Susan C., Kaleko, Michael, Smith, Theodore and
Nemerow, Glen R. Where permitted, the subject matter of each of these
applications is incorporated by reference herein.

15 FIELD OF INVENTION

The present invention generally relates to the field of adenoviral vectors and the production of such vectors. In particular, detargeted adenoviral vectors are provided.

BACKGROUND

20 Most, if not all, adenoviral vector-mediated gene therapy strategies aim to transduce a specific tissue, such as a tumor or an organ. Systemic delivery will require ablation of the normal virus tropism as well as addition of new specificities. Multiple interactions between adenoviral particles and the host cell are required to promote efficient cell entry (Nemerow (2000) *Virology* 274:1-4).
25 An adenovirus entry pathway is believed to involve two separate cell surface events. First, a high affinity interaction between the adenoviral fiber knob and coxsackie-adenovirus receptor (CAR) mediates the attachment of the adenovirus particle to the cell surface. A subsequent association of penton with the cell surface integrins $\alpha_v\beta_3$ and $\alpha_v\beta_5$, which act as co-receptors, potentiates virus
30 internalization. There are a plurality of adenoviral fiber receptors, which interact with the group B (e.g., Ad3) and group C (e.g., Ad5) adenoviruses. Both of these groups of adenoviruses appear to require interaction with integrins for

-2-

internalization. CAR ablation, however, does not change biodistribution and toxicity of adenoviral vectors *in vivo* (Alemany *et al.* (2001) *Gene Therapy* 8:1347-1353; U.S. patent application No. 09/870,203, filed May 30, 2001, and published as U.S. Published application No. 20020137213). Thus, the role of CAR interaction for *in vivo* gene transfer is not clear. Recently published studies have described conflicting results (Alemany *et al.* (2001) *Gene Therapy* 8:1347-1353; Leissner *et al.* (2001) *Gene Therapy* 8:49-57; Einfeld *et al.* (2001) *J. Virology* 75:11284-11291). For example, it has been shown that vectors containing an S408E mutation in the Ad5 fiber AB loop yield efficient liver transduction in mice, despite having greatly reduced transduction efficiencies on cells in culture (see, Leissner *et al.* (2001) *Gene Therapy* 8:49-57). In contrast, vectors containing a more extensive fiber AB loop mutation showed a 10-fold reduction in liver gene expression (see, Einfeld *et al.* (2001) *J. Virology* 75:11284-11291).

A doubly ablated adenovirus has been prepared by modifying the CAR binding region in the fiber loop and the integrin binding region in the penton base (Einfeld *et al.* (2001) *J. Virology* 75:11284-11291). This doubly ablated adenovirus, lacking CAR and integrin interactions, was reported not only to lack *in vitro* transduction of various cell types but also to lack *in vivo* transduction of liver cells. Specifically, the doubly ablated adenovirus was reported to have a 700 fold reduction in liver transduction when compared to the non-ablated adenovirus. These results, however, were not reproduced by others.

For many applications, the most clinically useful adenoviral vector would be deliverable systemically, such as into a peripheral vein, and would be targeted to a desired location in the body, and would not have undesirable side effects resulting from targeting to other locations. *In vivo* adenoviral vector targeting is a major goal in gene therapy and a significant effort has been focused on developing strategies to achieve this goal. Successful targeting strategies would direct the entire vector dose to the appropriate site and would be likely to improve the safety profile of the vector by permitting the use of lower, less toxic vector doses, which potentially also can be less immunogenic. Thus, there is a

-3-

need to develop adenoviruses which are fully detargeted *in vivo* for use as a base vector for producing redirected adenoviruses.

Therefore, among the objects herein, it is an object herein to provide fully detargeted adenoviral vectors, methods for preparation thereof, and uses

5 thereof.

SUMMARY

Detargeted and fully detargeted adenoviral particles, adenovirus vectors from which such particles are produced, methods for preparation of the vectors and particles and uses of the vectors and particles are provided. Provided and
10 described are capsid modifications, such as fiber shaft modifications, and the resulting proteins that, when expressed on adenoviral particles provide for detargeting of adenoviral vectors. The capsid modifications, such as the fiber shaft modifications, can be combined with other modifications, such as fiber knob and/or penton modifications, to produce fully ablated (detargeted)
15 adenoviral particles. Thus, adenoviral vectors and adenoviral particles whose native tropisms are ablated through a modification or modifications of capsid proteins, particularly a fiber shaft region, are provided.

Thus, provided are capsid mutations, including fiber shaft modifications, that ablate binding to particular receptors, thereby permitting efficient targeting
20 of adenoviral vectors that contain capsids with such modifications. For example, adenoviral vectors in which the fiber shaft's interaction with HSP is ablated (reduced or substantially eliminated), particularly *in vivo*, are provided. These fiber shaft modifications can be combined with other modifications, such as fiber knob and/or penton modifications, to produce fully ablated (detargeted)
25 adenoviral vectors. Also provided are retargeted vectors and particles that include a ligand or ligands to provide for targeting of the detargeted vectors and particles to selected cells and/or tissues. Retargeting can be effected, for example, by manipulating the fiber protein to redirect the receptor specificity to a particular cell type.

30 Also provided are nucleic acids encoding the modified fiber proteins and also modified penton proteins. Also provided are nucleic acids encoding the modified fiber shaft protein that has ablated HSP binding and combinations

-4-

thereof with other modified fiber regions or other proteins, such as a modified fiber knob region and/or the modified penton protein. The nucleic acids also can contain heterologous nucleic acid sequences, such as promoters or nucleic acid sequences encoding polypeptides. The viral particles that express fibers
5 containing such shaft modifications and other modifications are also provided.

Also provided are methods for making and using the adenoviral particles that express the modified fibers and combinations of modified fibers and modified penton. With the fiber shaft modifications, particularly in combination with the fiber knob modifications and the penton modifications, the adenovirus
10 particles are ablated for binding to their natural cellular receptor(s), *i.e.*, they are detargeted. They can then be "retargeted" to a specific cell type through the addition of a ligand to the virus capsid, which causes the virus to bind to and infect such cell. The ligand can be added, for example, through genetic modification of a capsid protein gene.

Also provided is a method for reducing liver toxicity in adeno-
15 viral-mediated therapy. In contrast to the results of Einfeld *et al.* (Einfeld *et al.* (2001) *J. Virology* 75:11284-11291), it is shown herein that a doubly ablated adenovirus, lacking CAR and integrin interactions, is capable of *in vivo* liver transduction. It is shown herein that ablation of liver transduction requires
20 further and/or alternative modification(s). The method for reducing liver toxicity in adenoviral-mediated therapy includes modifying an adenoviral vector to ablate native tropism to liver cells *in vivo*. Such vector can be administered to a subject. The modifications include the modifications described herein.

25 The nucleic acids, proteins, adenoviral particles and adenoviral vectors have a variety of uses. These include *in vivo* and *in vitro* uses to target nucleic acid to particular cells and tissues, for therapeutic purposes, including gene therapy, and also for the identification and study of cell surface receptors and identification of modes of interaction of viruses with cells.

30 In particular, adenoviral fiber shaft modifications that ablate viral interaction with HSP (Heparin Sulfate Proteoglycans; also referred to as heparin sulfate glycosaminoglycans) are provided. These modifications include

-5-

mutations of individual amino acids in the fiber shaft that interact with HSP or mutations of amino acids in the fiber shaft that modify the ability of the HSP binding motif to interact with HSP. Adenoviral fiber shaft modifications also include replacements of fiber shafts using fiber shafts of adenoviruses, such as, for example, Ad3, Ad35 and Ad41 short fiber shaft, that do not contain HSP binding sites.

Also provided are adenoviral fiber shaft modifications that alter, particularly ablate viral interaction with HSP, as described above, in combination with fiber knob modifications that ablate viral interaction with CAR. The fiber knob modifications include: (a) mutations of individual amino acids in the fiber loop that interact with CAR, such as, for example, AB or CD loop modifications; (b) mutations of individual amino acids in the fiber loop that modify the ability of the CAR binding motif to interact with CAR; and (c) replacements of fiber knobs using adenoviruses that do not interact with CAR, such as, for example, Ad3 fiber knob, Ad41 short fiber knob, or Ad35 fiber knob.

Also provided are adenoviral fiber shaft modifications as described above in combination with penton modifications that ablate viral interaction with α_v integrins. The penton modifications include: (a) mutations of individual amino acids that interact with α_v integrins; (b) mutations of individual amino acids that modify the ability of the α_v integrin binding motif to interact with the α_v integrins; and (c) replacement of penton proteins using penton proteins from adenoviruses that do not interact with the α_v integrins.

Also provided are adenoviral fiber shaft modifications as described above in combination with fiber knob modifications as described above and penton modifications as described above.

Also provided is a scale-up method for the propagation of detargeted adenoviral vectors. The method uses polycations and/or bifunctional reagents, which when added to tissue culture medium results in entry of adenoviral particles into the producer cells.

-6-

Provided are recombinant viral particles that contain a modified capsid protein whereby binding to heparin sulfate proteoglycans (HSP) is reduced or eliminated compared to particles that contain unmodified capsid proteins. The modified capsid proteins include fiber proteins with modified shafts such that
5 binding to HSP is reduced or eliminated.

Among the particular embodiments the following are provided. Provided are adenovirus capsid proteins that are modified to alter, typically reduce or eliminate, binding to or interaction with *in vivo* and/or *in vitro* to heparin sulfate proteoglycan (HSP). HSPs are expressed on various cells, including
10 hepatocytes. It is shown herein that HSPs provide for or participate in transduction of cells, such as liver cells. Since it can be desirable to eliminate or reduce such transduction, the modifications of the capsid proteins, such as fiber proteins, permit detargeting of particles that express such proteins from such cells.

15 Thus provided are modified adenovirus fiber proteins that include a mutation, such as an insertion, deletion, change, replacement of amino acids or combinations thereof, whereby binding to or interaction with heparin sulfate proteoglycan (HSP) is altered. In particular, the binding of the modified fiber protein is eliminated or reduced compared to the unmodified protein. Exemplary
20 of these mutations are mutations in the shaft of a fiber, where the shaft also can include the tail. The mutations can reduce or alter the affinity of the fiber protein for HSP is reduced at least by 2-fold, 5-fold, 10-fold, 100-fold or more, including substantially eliminating it.

As provided herein, fibers from adenoviruses that interact with HSP can
25 include a motif, such as BBXB or BBBXXB, where the B is a basic amino acid and X is any amino acid, particularly the consensus sequence KKTK in Ad5 and Ad2. Thus, provided are fibers in which the motif is altered to eliminate or reduce interaction with HSP.

Also provided are modified fiber protein of claim 1 that are chimeras in
30 which the fiber shaft (or fiber shaft and tail) are derived from a fiber, such as Ad3, Ad35, Ad7, Ad11, Ad16, Ad21, Ad34, Ad40, Ad41 or Ad46 fiber, that does not interact with HSP and combined with fiber that does interact, such as

-7-

Ad5 or Ad2 fiber, to produce a complete fiber whose binding to HSP is reduced or eliminated.

All of the modified capsid proteins provided herein also can include one or more further modifications that reduce or eliminate interaction of the resulting
5 fiber with one or more cell surface proteins, such as but not limited to, CAR and α_v integrin or other receptor to which a particular native fiber binds, in addition to HSP. These modifications include, but are not limited to, modification to fiber that reduces or eliminates CAR binding and modification to penton that reduces or eliminate α_v integrin binding. The mutations can be in the fiber knob, shaft,
10 tail and shaft, and also in penton.

Any and all of the modified capsid proteins provided herein can further include a ligand that binds to a particular receptor thereby endowing a fiber (or other capsid protein) with binding specificity or the ability to interact with such receptor. The ligand can be inserted into any suitable site in a capsid protein,
15 such as an insertion or replacement. For example, fibers with ligands inserted into the knob region are exemplified. Any such ligand can be employed and a variety are exemplified herein.

A variety of modified capsid proteins are exemplified herein. These include, but are not limited to, fibers containing: the sequence of amino acids
20 set forth in any of SEQ ID Nos. 52, 54, 56, 58, 62, 66, 70 and 72; or a sequence of amino acids having 60%, 70%, 80%, 90%, 95% or greater sequence identity with a sequence of amino acids set forth in any of SEQ ID Nos. 52, 54, 56, 58, 62, 66, 70 and 72; or a sequence of amino acids encoded by a sequence of nucleotides that hybridizes under conditions of high stringency
25 along at least 70% of its length to a sequence of nucleotides that encodes a sequence of amino acids set forth in any of SEQ ID Nos. 52, 54, 56, 58, 62, 66, 70 and 72.

Nucleic acids encoding the capsid proteins, including the fibers are also provided. The nucleic acids can be provided as vectors, particularly as
30 adenovirus vectors. Many adenoviral vectors are known and can be modified as needed in accord with the description herein. Adenoviral vectors include, but are not limited to, early generation adenoviral vectors, such as E1-deleted

-8-

vectors, gutless adenoviral vectors and replication-conditional adenoviral vectors, such as oncolytic adenoviral vectors. The adenovirus vectors also can include heterologous nucleic acids that encode or provide products, such as therapeutic products. Any therapeutic product is contemplated and a variety are set forth

- 5 herein as exemplary. Heterologous nucleic acid can encode a polypeptide or comprise or encode a regulatory sequence, such as a promoter or an RNA, including RNAi, small RNAs, other double-stranded RNAs, antisense RNA, and ribozymes. Promoters include, for example, constitutive and regulated promoters and tissue specific promoter, including tumor specific promoters.
- 10 The promoter can be operably linked, for example, to a gene of an adenovirus essential for replication.

Cells containing the nucleic acid molecules and cells containing the vectors are also provided. Such cell include packaging cells. The cells can be prokaryotic or eukaryotic cells, including, mammalian cells, such a primate cells,

15 including human cells.

- Also provided are adenoviral particles that contain the modified capsid proteins provided herein. The particles have altered interaction or binding with HSP compared to particles that do not contain the modified capsid proteins. In addition to altered binding to HSP, which is typically reduced or eliminated
- 20 binding, the particles can include further modifications, such as capsid proteins with altered interaction with other receptors as described above. In particular, the particles can have altered, typically reduced or eliminated, interaction with CAR, α_v integrin and/or other receptors. The mutation include mutations in the fiber knob, penton and hexon. Exemplary fiber knob mutations are mutations in
- 25 the AB loop or CD loop, such as KO1 or KO12, which are described herein. In addition, the particles can include additional ligands for retargeting to selected receptors. The adenoviral particles can be from any serotype and subgroup.

- Methods for expressing heterologous nucleic acids in a cell are provided. In these methods an adenoviral vector provided herein is transduced into a cell
- 30 to deliver the nucleic acid and/or encoded products. Transduction can be effected *in vivo* or *in vitro* or *ex vivo*, and can be for a variety of purposes including study of gene expression and genetic therapy. The cells can be

-9-

prokaryotic cells, but typically are eukaryotic cells, including mammalian cells, such as primate, including human, cells. The cells can be of a specific type, such as a tumor cell or a cell in a particular tissue. The vectors can be oncolytic vector to effect killing of tumor cells.

- 5 Since the modified capsid proteins herein have reduced or eliminated binding to HSP, viral particles containing such proteins exhibit ablated binding to HSP *in vitro* and *in vivo*. Thus provided is a method of reducing transduction of cells that express HSP, such as hepatocytes in the liver, by modifying a capsid protein, such as fiber to eliminate or reduce interaction with or binding to HSP.
- 10 Such reduction reduces or eliminates transduction of cells that express HSP, including liver cells.

- Also provided are scale-up methods for the propagation of detargeted adenoviral particle, such as those provided herein. The method includes the steps of infecting or transducing a cell capable of replicating, maturing and
- 15 packaging an adenoviral vector with a detargeted adenoviral vector in the presence of a reagent that results in entry of the adenoviral particle into the cell, such as a polycation and/or a bifunctional protein or other such reagent; and culturing the infected cell under conditions suitable for growth, spread and propagation of the adenoviral vector. The resulting adenoviral particles can be
- 20 recovered. Polycations include, but are not limited to, hexadimethrine bromide, polyethylenimine, protamine sulfate and poly-L-lysine. Bifunctional proteins, include, but are not limited to, an anti-fiber antibody ligand fusion, an anti-fiber-Fab-FGF conjugate, an anti-penton-antibody ligand fusion, an anti-hexon antibody ligand fusion and a polylysine-peptide fusion. The ligand is
- 25 selected to bind to a particular receptor.

- The viral particles that express a modified capsids provided herein can be produced by this method. The modification include, for example, one or more mutations selected from among mutations that reduce or eliminate interactions with one or more of α_v integrins, coxsackie-adenovirus receptors (CAR) and
- 30 heparin sulfate proteoglycans (HSP). Such mutations include, for example, PD1, KO1, KO12 and S*.

-10-

BRIEF DESCRIPTION OF THE FIGURES

Figure 1 is a plasmid map for pSKO1.

Figure 2 is a plasmid map for pNDSQ3.1KO1.

Figures 3A-3C are plasmid maps of pAdmireRSVnBg.(Fig. 3A), pSQ1 Fig.
5 3B) and pSQ1KO12 (Fig. 3C)

Figure 4 is a plasmid map for pSQ1PD1.

Figures 5A-5B are plasmid maps of pSQ1FKO1PD1 (Fig. 5A) and
pSQ1KO12PD1 (Fig. 5B).

Figure 6 shows *in vitro* transduction efficiency of A549 cells using
10 adenoviral vectors containing fiber AB loop knob and/or penton, PD1 mutations.
The following adenoviral vectors were used in these studies: Av1nBg,
Av1nBgFKO1, referred to as FKO1, Av1nBgPD1, referred to as PD1, and
Av1nBgFKO1PD1 that is referred to as FKO1PD1.

Figure 7A-7B shows *in vivo* adenoviral-mediated liver gene expression
15 (Fig. 7A) and hexon DNA content (Fig. 7B) using adenoviral vectors containing
fiber AB loop knob and/or penton, PD1 mutations. The following adenoviral
vectors were used in these studies: Av1nBg, Av1nBgFKO1, referred to as
FKO1, Av1nBgPD1, referred to as PD1, Av1nBgFKO1PD1, referred to as
FKO1PD1, Av1nBgKO12, referred to as KO12, and Av1nBgKO12PD1 that is
20 referred to as KO12PD1.

Figure 8 is a plasmid map for pFBshuttle(EcoRI).

Figure 9 is a plasmid map for pSQ1HSP.

Figure 10 is a plasmid map for pSQ1HSPKO1.

Figure 11 is a plasmid map for pSQ1HSPPD1.

25 Figure 12 is a plasmid map for pSQ1HSPKO1PD1.

Figures 13A-13C show the transduction efficiency of A549 and HeLa
cells using adenoviral vectors containing fiber shaft, knob and/or penton
mutations. Fig. 13A shows the dose response for the transduction efficiency of
A549 cells. Fig. 13B shows the transduction efficiency of HeLa cells at 2000
30 ppc. Figure 13C shows the competition analysis of adenoviral vectors
containing fiber shaft mutations.

-11-

Figures 14A-14B shows the influence of fiber shaft mutations on *in vivo* adenoviral-mediated liver gene expression (Fig. 14A) and hexon DNA content (Fig. 14B).

Figures 15A-15B are plasmid maps of pSQ1HSPRGD (Fig. 15A) and
5 pSQ1HSPKO1RGD (Fig. 15B).

Figure 16 shows that insertion of a RGD targeting ligand can restore transduction of the vectors containing the HSP binding shaft S* mutation.

Figures 17A-17B are plasmid maps of pSQ1AD35Fiber (Fig. 17A) and pSQ1Ad35FcRGD (Fig. 17B).

10 Figures 18A-18B are maps of plasmids encoding 35F chimeric fibers. Fig. 18A is a plasmid map of pSQ135T5H, and Fig. 18B is a plasmid map of pSQ15T35H.

Figure 19 shows the results of an *in vitro* analysis of Ad5 vectors containing Ad35 fibers and derivatives thereof.

15 Figure 20 shows the results of an *in vivo* analysis of Ad5 vectors containing Ad35 fibers and derivatives thereof.

Figures 21A-21B are plasmid maps of pSQ1Ad41sF (Fig. 21A) and pSQ1Ad41sFRGD (Fig. 21B).

20 Figure 22 shows the results of an *in vivo* analysis of Ad5 vectors containing Ad41 short fiber.

Figure 23 shows the *in vitro* analysis of Ad5 based vectors containing the Ad41 short fiber which has been re-engineered to contain a cRGD ligand in the HI loop.

25 Figure 24 shows enhanced transduction of AE1-2a cells with the Av3nBgFKO1 detargeted adenoviral vector using hexadimethrine bromide (HB), protamine sulfate (PS) and poly-lysine-RGD (K14) or the anti-penton-TNF α bifunctional protein (α pen-TNF).

Figure 25 shows ablation of HSP interaction decreases adenoviral-mediated gene transfer to other organs

30 Figure 26 shows *in vivo* liver transduction with adenoviral vectors which encode for B-galactosidase and contain various mutations to the fiber and/or penton proteins. Results are plotted as percent transduction as compared to

-12-

wild type. Two different methods for determining the level of transduction are shown for each vector.

Figure 27 shows the adenoviral vector biodistribution to the liver and tumor for the vectors containing the S*, KO1S*, and 41sF fibers.

5 DETAILED DESCRIPTION

A. DEFINITIONS

Unless defined otherwise, all technical and scientific terms used herein have the same meaning as is commonly understood by one of skill in the art to which the invention(s) belong. All patents, patent applications, published applications and publications, Genbank sequences, websites and other published materials referred to throughout the entire disclosure herein, unless noted otherwise, are incorporated by reference in their entirety. In the event that there are a plurality of definitions for terms herein, those in this section prevail.

Where reference is made to a URL or other such identifier or address, it is understood that such identifiers can change and particular information on the internet can come and go, but equivalent information is known and can be readily accessed, such as by searching the internet and/or appropriate databases. Reference thereto evidences the availability and public dissemination of such information.

As used herein, the term "adenovirus" or "adenoviral particle" is used to include any and all viruses that can be categorized as an adenovirus, including any adenovirus that infects a human or an animal, including all groups, subgroups, and serotypes. Depending upon the context reference to "adenovirus" can include adenoviral vectors. There are at least 51 serotypes of Adenovirus that classified into several subgroups. For example, subgroup A includes adenovirus serotypes 12, 18, and 31. Subgroup C includes adenovirus serotypes 1, 2, 5, and 6. Subgroup D includes adenovirus serotype 8, 9, 10, 13, 15, 17, 19, 20, 22-30, 32, 33, 36-39, and 42-49. Subgroup E includes adenovirus serotype 4. Subgroup F includes adenovirus serotypes 40 and 41.

These latter two serotypes have a long and a short fiber protein. Thus, as used herein an adenovirus or adenovirus particle is a packaged vector or genome.

-13-

As used herein, "virus," "viral particle," "vector particle," "viral vector particle," and "virion" are used interchangeably to refer to infectious viral particles that are formed when, such as when a vector containing all or a part of a viral genome, is transduced into an appropriate cell or cell line for the
5 generation of such particles. The resulting viral particles have a variety of uses, including, but not limited to, transferring nucleic acids into cells either *in vitro* or *in vivo*. For purposes herein, the viruses are adenoviruses, including recombinant adenoviruses formed when an adenovirus vector, such as any provided herein, is encapsulated in an adenovirus capsid. Thus, a viral particle is a packaged viral
10 genome. An adenovirus viral particle is the minimal structural or functional unit of a virus. A virus can refer to a single particle, a stock of particles or a viral genome. The adenovirus (Ad) particle is relatively complex and may be resolved into various substructures.

Included among adenoviruses and adenoviral particles are any and all
15 viruses that can be categorized as an adenovirus, including any adenovirus that infects a human or an animal, including all groups, subgroups, and serotypes. Thus, as used herein, "adenovirus" and "adenovirus particle" refer to the virus itself and derivatives thereof and cover all serotypes and subtypes and naturally occurring and recombinant forms, except where indicated otherwise. Included
20 are adenoviruses that infect human cells. Adenoviruses can be wildtype or can be modified in various ways known in the art or as disclosed herein. Such modifications include, but are not limited to, modifications to the adenovirus genome that is packaged in the particle in order to make an infectious virus. Exemplary modifications include deletions known in the art, such as deletions in
25 one or more of the E1a, E1b, E2a, E2b, E3, or E4 coding regions. Other exemplary modifications include deletions of all of the coding regions of the adenoviral genome. Such adenoviruses are known as "gutless" adenoviruses. The terms also include replication-conditional adenoviruses, which are viruses that preferentially replicate in certain types of cells or tissues but to a lesser
30 degree or not at all in other types. For example, among the adenoviral particles provided herein, are adenoviral particles that replicate in abnormally proliferating tissue, such as solid tumors and other neoplasms. These include the viruses

-14-

disclosed in U.S. Patent No. 5,998,205 and U.S. Patent No. 5,801,029. Such viruses are sometimes referred to as "cytolytic" or "cytopathic" viruses (or vectors), and, if they have such an effect on neoplastic cells, are referred to as "oncolytic" viruses (or vectors).

5 As used herein, the terms "vector," "polynucleotide vector," "polynucleotide vector construct," "nucleic acid vector construct," and "vector construct" are used interchangeably herein to mean any nucleic acid construct that can be used for gene transfer, as understood by those skilled in the art.

 As used herein, the term "viral vector" is used according to its
10 art-recognized meaning. It refers to a nucleic acid vector construct that includes at least one element of viral origin and can be packaged into a viral vector particle. The viral vector particles can be used for the purpose of transferring DNA, RNA or other nucleic acids into cells either in vitro or in vivo. Viral vectors include, but are not limited to, retroviral vectors, vaccinia vectors, lentiviral
15 vectors, herpes virus vectors (e.g., HSV), baculoviral vectors, cytomegalovirus (CMV) vectors, papillomavirus vectors, simian virus (SV40) vectors, Sindbis vectors, semliki forest virus vectors, phage vectors, adenoviral vectors, and adeno-associated viral (AAV) vectors. Suitable viral vectors are described, for example, in U.S. Patent Nos. 6,057,155, 5,543,328 and 5,756,086. The
20 vectors provided herein are adenoviral vectors.

 As used herein, "adenovirus vector" and "adenoviral vector" are used interchangeably and are well understood in the art to mean a polynucleotide containing all or a portion of an adenovirus genome. An adenoviral vector, refers to nucleic encoding a complete genome or a modified genome or one that can be
25 used to introduce heterologous nucleic acid when transferred into a cell, particularly when packaged as a particle. An adenoviral vector can be in any of several forms, including, but not limited to, naked DNA, DNA encapsulated in an adenovirus capsid, DNA packaged in another viral or viral-like form (such as herpes simplex, and AAV), DNA encapsulated in liposomes, DNA complexed
30 with polylysine, complexed with synthetic polycationic molecules, conjugated with transferrin, complexed with compounds such as PEG to immunologically

-15-

"mask" the molecule and/or increase half-life, or conjugated to a non-viral protein.

As used herein, oncolytic adenoviruses refer to adenoviruses that replicate selectively in tumor cells

5 As used herein, a variety of vectors with different requirements and purposes are described. For example, one vector is used to deliver particular nucleic acid molecules into a packaging cell line for stable integration into a chromosome. These types of vectors also are referred to as complementing plasmids. A further type of vector carries or delivers nucleic acid molecules in or
10 into a cell line (*e.g.*, a packaging cell line) for the purpose of propagating viral vectors; hence, these vectors also can be referred to herein as delivery plasmids. A third "type" of vector is the vector that is in the form of a virus particle encapsulating a viral nucleic acid and that is comprised of the capsid modified as provided herein. Such vectors also can contain heterologous nucleic acid
15 molecules encoding particular polypeptides, such as therapeutic polypeptides or regulatory proteins or regulatory sequences to specific cells or cell types in a subject in need of treatment.

As used herein, the term "motif" is used to refer to any set of amino acids forming part of a primary sequence of a protein, either contiguous or
20 capable of being aligned to certain positions that are invariant or conserved, that is associated with a particular function. The motif can occur, not only by virtue of the primary sequence, but also as a consequence of three-dimensional folding. For example, the motif GXGXXG is associated with nucleotide-binding sites. In this fiber is a trimer, hence the trimeric structure can contribute formation of a
25 motif. Alternatively, a motif can be considered as a domain of a protein, where domain is a region of a protein molecule delimited on the basis of function without knowledge of and relation to the molecular substructure, as, *e.g.*, the part of a protein molecule that binds to a receptor. As shown herein, the motif KKTK constitutes a consensus sequence for fiber shaft interaction with HSP.

30 As used herein, the term "bind" or "binding" is used to refer to the binding between a ligand and its receptor, such as the binding of an Ad5 shaft motif with HSP (Heparin Sulfate Proteoglycans), with a K_d in the range of 10-2

-16-

to 10⁻¹⁵ mole/l, generally, 10⁻⁶ to 10⁻¹⁵, 10⁻⁷ to 10⁻¹⁵ and typically 10⁻⁸ to 10⁻¹⁵ (and/or a K_a of 10⁵-10¹², 10⁷-10¹², 10⁸-10¹² l/mole).

As used herein, specific binding or selective binding means that a the binding of a particular ligand and one receptor interaction (k_a or K_{eq}) is at least 2-

5 fold, generally, 5, 10, 50, 100 or more-fold, greater than for another receptor.

A statement that a particular viral vector is targeted to a cell or tissue means that its affinity for such cell or tissue in a host or *in vitro* is at least about 2-fold, generally, 5, 10, 50, 100 or more-fold, greater than for other cells and tissues in the host or under the *in vitro* conditions.

10 As used herein, the term "ablate" or "ablated" is used to refer to an adenovirus, adenoviral vector or adenoviral particle, in which the ability to bind to a particular cellular receptor is reduced or eliminated, generally substantially eliminated (*i.e.*, reduced more than 10-fold, 100-fold or more) when compared to a corresponding wild-type adenovirus. An ablated adenovirus, adenoviral vector
15 or adenoviral particle also is said to be detargeted, *i.e.*, the modified adenovirus, adenoviral vector or adenoviral particle does not possess the native tropism of the wild-type adenovirus. The reduction or elimination of the ability of the mutated adenovirus fiber protein and/or mutated adenovirus penton protein to bind a cellular receptor as compared to the corresponding wild-type fiber protein
20 and/or wild-type penton protein can be measured or assessed by comparing the transduction efficiency (gene transfer and expression of a marker gene) of an adenovirus particle containing the mutated fiber protein and/or mutated penton protein compared to an adenovirus particle containing the wild-type fiber protein and/or wild-type penton protein for cells having the cellular receptor.

25 As used herein, tropism with reference to an adenovirus refers to the selective infectivity or binding that is conferred on the particle by a capsid protein, such as the fiber protein and/or penton.

As used herein, "penton" or "penton complex" is used herein to designate a complex of penton base and fiber. The term "penton" can also be used to
30 indicate penton base, as well as penton complex. The meaning of the term "penton" alone should be clear from the context within which it is used.

-17-

As used herein, the term "substantially eliminated" refers to a transduction efficiency less than about 11% of the efficiency of the wild-type fiber containing virus on HeLa cells. The transduction efficiency on HeLa cells can be measured (see, *e.g.*, Example 1 of U.S. Patent Application Serial No. 09/870,203 filed on 30 May 2001, and published as U.S. Published application No. 20020137213, and of International Patent Application No. PCT/EP01/06286 filed 1 June 2001). Briefly, HeLa cells are infected with the adenoviral vectors containing mutated fiber proteins to evaluate the effects of fiber amino acid mutations on CAR interaction and subsequent gene expression. Monolayers of HeLa cells in 12 well dishes are infected with, for example, 1000 particles per cell for 2 hours at 37° C. in a total volume of, for example, 0.35 ml of the DMEM containing 2% FBS. The infection medium is then aspirated from the monolayers and 1 ml of complete DMEM containing 10% FBS was added per well. The cells are incubated for an sufficient time, generally about 24 hours, to allow for β -galactosidase expression, which is measured by a chemiluminescence reporter assay and by histochemical staining with a chromogenic substrate. The relative levels of β -galactosidase activity are determined using as suitable system, such as the Galacto-Light chemiluminescence reporter assay system (Tropix, Bedford, Mass.) Cell monolayers are washed with PBS and processed according to the manufacturer's protocol. The cell homogenate is transferred to a microfuge tube and centrifuged to remove cellular debris. Total protein concentration is determined, such as by using the bicinchoninic acid(BCA) protein assay (Pierce, Inc., Rockford, Ill.) with bovine serum albumin as the assay standard. An aliquot of each sample is then incubated with the Tropix β -galactosidase substrate for 45 minutes in a 96 well plate. A luminometer is used determine the relative light units (RLU) emitted per sample and then normalized for the amount of total protein in each sample (RLU/ μ g total protein). For the histochemical staining procedure, the cell monolayers are fixed with 0.5% glutaraldehyde in PBS, and then were incubated with a mixture of 1 mg of 5-bromo-4-chloro-3-indolyl- β -D-galactoside (X-gal) per ml, 5 mM potassium ferrocyanide, 5 mM potassium ferricyanide and 2 mM $MgCl_2$ in 0.5 ml of PBS. The monolayers are washed with PBS and the blue cells are visualized by light

-18-

microscopy, such as with a Zeiss ID03 microscope. Generally, the efficiency is less than about 9%, and typically is less than about 8%.

As used herein, the phrase "reduce" or "reduction" refers to a change in the efficiency of transduction by the adenovirus containing the mutated fiber as compared to the adenovirus containing the wild-type fiber to a level of about 75% or less of the wild-type on HeLa cells. Generally, the change in efficiency is to a level of about 65% or less than wild-type. Typically it is about 55% or less. This system is able to rapidly analyze modified fiber proteins and/or modified penton proteins for desired tropism in the context of the viral particle.

As used herein, the term "mutate" or "mutation" or similar terms refers to the deletion, insertion or change of at least one amino acid in the part of the fiber shaft region interacting with HSP. The amino acid can be changed by substitution or by modification in a way that derivatizes the amino acid. Thus, for example a BBXB motif or BBBXXB motif, where B is a basic amino acid, in an adenovirus is mutated to ablate the viral interaction with HSP.

As used herein, the term "polynucleotide" means a nucleic acid molecule, such as DNA or RNA, that encodes a polynucleotide. The molecule can include regulatory sequences, and is generally DNA. Such polynucleotides are prepared or obtained by techniques known by those skilled in the art in combination with the teachings contained therein.

As used herein, adenoviral genome is intended to include any adenoviral vector or any nucleic acid sequence comprising a modified fiber protein. All adenovirus serotypes are contemplated for use in the vectors and methods herein.

As used herein, the term "viral vector" is used according to its art-recognized meaning. It refers to a nucleic acid vector construct that includes at least one element of viral origin and can be packaged into a viral vector particle. The viral vector particles can be used, for example, for transferring DNA into cells either *in vitro* or *in vivo*.

As used herein, a packaging cell line is a cell line that is able to package adenoviral genomes or modified genomes to produce viral particles. It can provide a missing gene product or its equivalent. Thus, packaging cells can

-19-

provide complementing functions for the genes deleted in an adenoviral genome (*e.g.*, the nucleic acids encoding modified fiber proteins) and are able to package the adenoviral genomes into the adenovirus particle. The production of such particles require that the genome be replicated and that those proteins necessary
5 for assembling an infectious virus are produced. The particles also can require certain proteins necessary for the maturation of the viral particle. Such proteins can be provided by the vector or by the packaging cell.

As used herein, detargeted adenoviral particles have ablated (reduced or eliminated) interaction with receptors with which native particles. Fully
10 detargeted particles have two or more specificities altered. It is understood that *in vivo* no particles are fully ablated such that they do not interact with any cells. Degareted and fully degareted have reduced, typically substantiall reduced, or eliminated interaction with native receptors. For purposes herein, detargeted particles have reduced (2-fold, 5-fold, 10-fold, 100-fold or more) binding or
15 virtually no binding to HSP receptors; fully degareted vectors include further capsid modifications to eliminate interactions with other receptors, such as CAR and integrins or other receptors. The particles still bind to cells, but the types of cells and interactions are reduced.

As used herein, pseudotyping describes the production of adenoviral
20 vectors having modified capsid protein or capsid proteins from a different serotype than the serotype of the vector itself. One example, is the production of an adenovirus 5 vector particle containing an Ad37 or Ad35 fiber protein. This can be accomplished by producing the adenoviral vector in packaging cell lines expressing different fiber proteins. As provided herein, detargeting of an
25 adenovirus 5 particle or other serotype group C adenovirus or other adenovirus that binds to HSP to reduce or eliminate binding to HSPs can be effected by replacing all or a portion that includes the shaft or at least the HSP consensus binding sequence of the Ad5 fiber with an adenovirus fiber or portion thereof that does not bind to HSP. Adenoviruses having fiber shafts that do not interact
30 with HSP include (a) adenoviruses of subgroup B, *e.g.*, Ad3, Ad35, Ad7, Ad11, Ad16, Ad21, Ad34 which do not have interaction with HSP, (b) adenoviruses of

-20-

subgroup F, e.g., Ad40 and Ad41, specifically the short fiber, and (c) adenoviruses of subgroup D, e.g., Ad46.

As used herein, receptor refers to a biologically active molecule that specifically or selectively binds to (or with) other molecules. The term "receptor protein" can be used to more specifically indicate the proteinaceous nature of a
5 specific receptor.

As used herein, the term "cyclic RGD" (or cRGD) refers to any amino acid that binds to α_v integrins on the surface of cells and contains the sequence RGD (Arg-Gly-Asp).

10 As used herein, the term "heterologous polynucleotide" means a polynucleotide derived from a biological source other than an adenovirus or from an adenovirus of a different strain or can be a polynucleotide that is in a different locus from wild-type virus. The heterologous polynucleotide can encode a polypeptide, such as a toxin or a therapeutic protein. The heterologous
15 polynucleotide can contain regulatory regions, such as a promoter regions, such as a promoter active in specific cells or tissue, for example, tumor tissue as found in oncolytic adenoviruses. Alternatively, the heterologous polynucleotide can encode a polypeptide and further contain a promoter region operably linked to the coding region.

20 As used herein, reference to an amino acid in an adenovirus protein or to a nucleotide in an adenovirus genome is with reference to Ad5, unless specified otherwise. Corresponding amino acids and nucleotides in other adenovirus strains and modified strains and in vectors can be identified by those of skill in the art. Thus recitation of a mutation is intended to encompass all adenovirus
25 strains that process a corresponding locus.

As used herein, the KO mutations refer to mutations in fiber that knock out binding to CAR. For example, a KO1 mutation refers to a mutation in the Ad5 fiber and corresponding mutations in other fiber proteins. In Ad5, this mutation results in a substitution of fiber amino acids 408 and 409, changing
30 them from serine and proline to glutamic acid and alanine, respectively. As used herein, a KO12 mutation refers to a mutation in the Ad5 fiber and corresponding mutations in other fiber proteins. In Ad5, this mutation is a four amino acid

-21-

substitution as follows: R512S, A515G, E516G, and K517G. Other KO mutations can be identified empirically or are known to those of skill in the art.

As used herein, PD mutations refer to mutations in the penton gene that ablate binding by the encoded α_v integrin by replacing the RGD tripeptide.

- 5 The PD1 mutation exemplified herein results in a substitution of amino acids 337 through 344 of the Ad5 penton protein, HAIRGDTF (SEQ ID No. 9), with amino acids SRGYPYDVPDYAGTS (SEQ ID No. 10), thereby replacing the RGD tripeptide.

- 10 As used herein, treatment means any manner in which the symptoms of a condition, disorder or disease are ameliorated or otherwise beneficially altered.

As used herein, a therapeutically effective product is a product that is encoded by heterologous DNA that, upon introduction of the DNA into a host, a product is expressed that effectively ameliorates or eliminates the symptoms, manifestations of an inherited or acquired disease or that cures said disease.

- 15 As used herein, a subject is an animal, such as a mammal, typically a human, including patients.

- 20 As used herein, genetic therapy involves the transfer of heterologous DNA to the certain cells, target cells, of a mammal, particularly a human, with a disorder or conditions for which such therapy is sought. The DNA is introduced into the selected target cells in a manner such that the heterologous DNA is expressed and a therapeutic product encoded thereby is produced.

- Alternatively, the heterologous DNA may in some manner mediate expression of DNA that encodes the therapeutic product, it may encode a product, such as a peptide or RNA that in some manner mediates, directly or indirectly, expression of a therapeutic product. Genetic therapy may also be used to deliver nucleic acid encoding a gene product to replace a defective gene or supplement a gene product produced by the mammal or the cell in which it is introduced. The introduced nucleic acid may encode a therapeutic compound, such as a growth factor inhibitor thereof, or a tumor necrosis factor or inhibitor thereof, such as a receptor therefor, that is not normally produced in the mammalian host or that is not produced in therapeutically effective amounts or at a therapeutically useful time. The heterologous DNA encoding the therapeutic product may be modified
- 25
- 30

-22-

prior to introduction into the cells of the afflicted host in order to enhance or otherwise alter the product or expression thereof.

As used herein, a therapeutic nucleic acid is a nucleic acid that ends a therapeutic product. The product can be nucleic acid, such as a regulatory
5 sequence or gene, or can encode a protein that has a therapeutic activity or effect. For example, therapeutic nucleic acid can be a ribozyme, antisense, double-stranded RNA, a nucleic acid encoding a protein and others.

As used herein, "homologous" means about greater than 25% nucleic acid sequence identity, such as 25%, 40%, 60%, 70%, 80%, 90% or 95%. If
10 necessary the percentage homology will be specified. The terms "homology" and "identity" are often used interchangeably. In general, sequences are aligned so that the highest order match is obtained (see, e.g.: *Computational Molecular Biology*, Lesk, A.M., ed., Oxford University Press, New York, 1988;

Biocomputing: Informatics and Genome Projects, Smith, D.W., ed., Academic
15 Press, New York, 1993; *Computer Analysis of Sequence Data, Part I*, Griffin, A.M., and Griffin, H.G., eds., Humana Press, New Jersey, 1994; *Sequence Analysis in Molecular Biology*, von Heinje, G., Academic Press, 1987; and *Sequence Analysis Primer*, Gribskov, M. and Devereux, J., eds., M Stockton Press, New York, 1991; Carillo *et al.* (1988) *SIAM J Applied Math* 48:1073).

20 By sequence identity, the number of conserved amino acids are determined by standard alignment algorithms programs, and are used with default gap penalties established by each supplier. Substantially homologous nucleic acid molecules would hybridize typically at moderate stringency or at high stringency all along the length of the nucleic acid or along at least about 70%, 80% or 90% of the
25 full-length nucleic acid molecule of interest. Also contemplated are nucleic acid molecules that contain degenerate codons in place of codons in the hybridizing nucleic acid molecule.

Whether any two nucleic acid molecules have nucleotide sequences that are at least, for example, 80%, 85%, 90%, 95%, 96%, 97%, 98% or 99%
30 "identical" can be determined using known computer algorithms such as the "FAST A" program, using for example, the default parameters as in Pearson *et al.* (1988) *Proc. Natl. Acad. Sci. USA* 85:2444 (other programs include the GCG

-23-

program package (Devereux, J., *et al.*, *Nucleic Acids Research* 12(1):387 (1984)), BLASTP, BLASTN, FASTA (Atschul, S.F., *et al.*, *J Molec Biol* 215:403 (1990); Guide to Huge Computers, Martin J. Bishop, ed., Academic Press, San Diego, 1994, and Carillo *et al.* (1988) *SIAM J Applied Math* 48:1073). For
5 example, the BLAST function of the National Center for Biotechnology Information database can be used to determine identity. Other commercially or publicly available programs include, DNASTar "MegAlign" program (Madison, WI) and the University of Wisconsin Genetics Computer Group (UWG) "Gap" program (Madison WI)). Percent homology or identity of proteins and/or nucleic
10 acid molecules can be determined, for example, by comparing sequence information using a GAP computer program (*e.g.*, Needleman *et al.* (1970) *J. Mol. Biol.* 48:443, as revised by Smith and Waterman ((1981) *Adv. Appl. Math.* 2:482). Briefly, the GAP program defines similarity as the number of aligned symbols (*i.e.*, nucleotides or amino acids) which are similar, divided by the total
15 number of symbols in the shorter of the two sequences. Default parameters for the GAP program can include: (1) a unary comparison matrix (containing a value of 1 for identities and 0 for non-identities) and the weighted comparison matrix of Gribskov *et al.* (1986) *Nucl. Acids Res.* 14:6745, as described by Schwartz and Dayhoff, eds., *ATLAS OF PROTEIN SEQUENCE AND STRUCTURE*, National
20 Biomedical Research Foundation, pp. 353-358 (1979); (2) a penalty of 3.0 for each gap and an additional 0.10 penalty for each symbol in each gap; and (3) no penalty for end gaps. Therefore, as used herein, the term "identity" represents a comparison between a test and a reference polypeptide or polynucleotide.

As used herein, the term "at least 90% identical to" refers to percent
25 identities from 90 to 99.99 relative to the reference polypeptides. Identity at a level of 90% or more is indicative of the fact that, assuming for exemplification purposes a test and reference polynucleotide length of 100 amino acids are compared, no more than 10% (*i.e.*, 10 out of 100) of amino acids in the test polypeptide differs from that of the reference polypeptides. Similar comparisons
30 can be made between a test and reference polynucleotides. Such differences can be represented as point mutations randomly distributed over the entire length of an amino acid sequence or they can be clustered in one or more

-24-

locations of varying length up to the maximum allowable, e.g. 10/100 amino acid difference (approximately 90% identity). Differences are defined as nucleic acid or amino acid substitutions, or deletions. At the level of homologies or identities above about 85-90%, the result should be independent of the program and gap parameters set; such high levels of identity can be assessed readily, often without relying on software.

As used herein: stringency of hybridization in determining percentage mismatch is as follows:

- 1) high stringency: 0.1 x SSPE, 0.1% SDS, 65°C
- 2) medium stringency: 0.2 x SSPE, 0.1% SDS, 50°C
- 3) low stringency: 1.0 x SSPE, 0.1% SDS, 50°C

Those of skill in this art know that the washing step selects for stable hybrids and also know the ingredients of SSPE (see, e.g., Sambrook, E.F. Fritsch, T. Maniatis, in: *Molecular Cloning, A Laboratory Manual*, Cold Spring Harbor Laboratory Press (1989), vol. 3, p. B.13, see, also, numerous catalogs that describe commonly used laboratory solutions). SSPE is pH 7.4 phosphate-buffered 0.18 M NaCl. Further, those of skill in the art recognize that the stability of hybrids is determined by T_m , which is a function of the sodium ion concentration and temperature ($T_m = 81.5^\circ \text{C} - 16.6(\log_{10}[\text{Na}^+]) + 0.41(\% \text{G} + \text{C}) - 600/l$), so that the only parameters in the wash conditions critical to hybrid stability are sodium ion concentration in the SSPE (or SSC) and temperature.

It is understood that equivalent stringencies can be achieved using alternative buffers, salts and temperatures. By way of example and not limitation, procedures using conditions of low stringency are as follows (see also Shilo and Weinberg, *Proc. Natl. Acad. Sci. USA* 78:6789-6792 (1981)): Filters containing DNA are pretreated for 6 hours at 40°C in a solution containing 35% formamide, 5X SSC, 50 mM Tris-HCl (pH 7.5), 5 mM EDTA, 0.1% PVP, 0.1% Ficoll, 1% BSA, and 500 µg/ml denatured salmon sperm DNA (10X SSC is 1.5 M sodium chloride, and 0.15 M sodium citrate, adjusted to a pH of 7).

Hybridizations are carried out in the same solution with the following modifications: 0.02% PVP, 0.02% Ficoll, 0.2% BSA, 100 µg/ml salmon sperm DNA, 10% (wt/vol) dextran sulfate, and 5-20 X 10⁶ cpm ³²P-labeled probe is

-25-

used. Filters are incubated in hybridization mixture for 18-20 hours at 40°C, and then washed for 1.5 hours at 55°C in a solution containing 2X SSC, 25 mM Tris-HCl (pH 7.4), 5 mM EDTA, and 0.1% SDS. The wash solution is replaced with fresh solution and incubated an additional 1.5 hours at 60°C. Filters are
5 blotted dry and exposed for autoradiography. If necessary, filters are washed for a third time at 65-68°C and reexposed to film. Other conditions of low stringency which can be used are well known in the art (*e.g.*, as employed for cross-species hybridizations).

By way of example and not way of limitation, procedures using
10 conditions of moderate stringency include, for example, but are not limited to, procedures using such conditions of moderate stringency are as follows: Filters containing DNA are pretreated for 6 hours at 55°C in a solution containing 6X SSC, 5X Denhart's solution, 0.5% SDS and 100 µg/ml denatured salmon sperm DNA. Hybridizations are carried out in the same solution and 5-20 X 10⁶ cpm
15 ³²P-labeled probe is used. Filters are incubated in hybridization mixture for 18-20 hours at 55°C, and then washed twice for 30 minutes at 60°C in a solution containing 1X SSC and 0.1% SDS. Filters are blotted dry and exposed for autoradiography. Other conditions of moderate stringency which can be used are well-known in the art. Washing of filters is done at 37°C for 1 hour in a
20 solution containing 2X SSC, 0.1% SDS.

By way of example and not way of limitation, procedures using conditions of high stringency are as follows: Prehybridization of filters containing DNA is carried out for 8 hours to overnight at 65°C in buffer composed of 6X SSC, 50 mM Tris-HCl (pH 7.5), 1 mM EDTA, 0.02% PVP, 0.02% Ficoll, 0.02% BSA,
25 and 500 µg/ml denatured salmon sperm DNA. Filters are hybridized for 48 hours at 65°C in prehybridization mixture containing 100 µg/ml denatured salmon sperm DNA and 5-20 X 10⁶ cpm of ³²P-labeled probe. Washing of filters is done at 37°C for 1 hour in a solution containing 2X SSC, 0.01% PVP, 0.01% Ficoll, and 0.01% BSA. This is followed by a wash in 0.1X SSC at 50°C for 45
30 minutes before autoradiography. Other conditions of high stringency which can be used are well known in the art.

-26-

The term substantially identical or substantially homologous or similar varies with the context as understood by those skilled in the relevant art and generally means at least 60% or 70%, preferably means at least 80%, 85% or more preferably at least 90%, and most preferably at least 95% identity.

- 5 As used herein, substantially identical to a product means sufficiently similar so that the property of interest is sufficiently unchanged so that the substantially identical product can be used in place of the product.

- 10 As used herein, substantially pure means sufficiently homogeneous to appear free of readily detectable impurities as determined by standard methods of analysis, such as thin layer chromatography (TLC), gel electrophoresis and high performance liquid chromatography (HPLC), used by those of skill in the art to assess such purity, or sufficiently pure such that further purification would not detectably alter the physical and chemical properties, such as enzymatic and biological activities, of the substance. Methods for purification of the
- 15 compounds to produce substantially chemically pure compounds are known to those of skill in the art. A substantially chemically pure compound can, however, be a mixture of stereoisomers or isomers. In such instances, further purification might increase the specific activity of the compound.

- 20 The methods and preparation of products provided herein, unless otherwise indicated, employ conventional techniques of chemistry, molecular biology, microbiology, recombinant DNA, genetics, immunology, cell biology, cell culture and transgenic biology, which are within the skill of the art (see, *e.g.*, Maniatis *et al.* (1982) *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY; Sambrook *et al.* (1989) *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY; Ausubel *et al.* (1992) *Current Protocols in Molecular Biology*, Wiley and Sons, New York; Glover (1985) *DNA Cloning I and II*, Oxford Press; Anand (1992) *Techniques for the Analysis of Complex Genomes* (Academic Press); Guthrie and Fink (1991) *Guide to Yeast Genetics and*
- 25 *Molecular Biology*, Academic Press; Harlow and Lane (1988) *Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY; Jakoby and Pastan, eds. (1979) *Cell Culture. Methods in Enzymology* 58,
- 30

-27-

Academic Press, Inc., Harcourt Brace Jovanovich, NY; *Nucleic Acid Hybridization* (B. D. Hames & S. J. Higgins eds. 1984); *Culture Of Animal Cells* (R. I. Freshney, Alan R. Liss, Inc., 1987); *Immobilized Cells And Enzymes* (IRL Press, 1986); B. Perbal (1984), *A Practical Guide To Molecular Cloning*; *Gene Transfer Vectors For Mammalian Cells* (J. H. Miller and M. P. Calos eds., 1987, Cold Spring Harbor Laboratory); *Methods In Enzymology*, Vols. 154 and 155 (Wu et al. eds.); *Immunochemical Methods In Cell And Molecular Biology* (Mayer and Walker, eds., Academic Press, London, 1987); *Handbook Of Experimental Immunology*, Volumes I-IV (D. M. Weir and C. C. Blackwell, eds., 1986); Hogan et al. (1986) *Manipulating the Mouse Embryo*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.

B. Capsid modifications

Provided herein are modifications of the viral capsid that ablate the interaction of an adenovirus with its natural receptors. In particular, fiber modifications that result in ablation of the interaction of an adenovirus with HSP are provided. These fiber modifications can be combined with other capsid protein modifications, such as other fiber modifications and/or penton and/or hexon modifications, to fully ablate viral interactions with natural receptors, when expressed on a viral particle. The modification should not disrupt trimer formation or transport of fiber into the nucleus.

1. Fiber genes and proteins

The fiber protein extends from the capsid and mediates viral binding to the cell surface by binding to specific cell receptors (Philipson et al. (1968) *J. Virol.* 2:1064-1075). The fiber is a trimeric protein that includes an N-terminal tail domain that interacts with the adenovirus penton base, a central shaft domain of varying length, and a C-terminal knob domain that contains the cell receptor binding site (Chroboczek et al. (1995) *Curr. Top. Microbiol. Immunol.* 199:163-200; Riurok et al. (1990) *J. Mol. Biol.* 215:589-596; Stevenson et al. (1995) *J. Virol.* 69:2850-2857). The sequences of the fiber gene from a variety of serotypes including adenovirus serotypes 2 (Ad2), Ad5, Ad3, Ad35, Ad12, Ad40, and Ad41 are known. There are at least 21 different fiber genes in Genbank.

-28-

As noted, the fiber protein can be divided into three domains (see, *e.g.*, Green *et al.* (1983) *EMBO J.* 2:1357-1365). The conserved N-terminus contains the sequences responsible for association with the penton base as well as a nuclear localization signal. A rod-like shaft of variable length contains repeats of a 15 amino acid beta structure, with the number of repeats ranging from 6 in Ad3 to 22 in Ad5. A conserved stretch of amino acids which includes the sequence TLWT (SEQ ID No. 36) marks the boundary between the repeating units of beta structure in the shaft and the globular head domain. The C-terminal head domain ranges in size from 157 amino acid residues for the short fiber of Ad41 to 193 residues for Ad11 and Ad34. The fiber spike is a homotrimer and it is thought that the C-terminus is responsible for trimerization of the fiber homotrimer and there are 12 spikes per virion which are attached via association with the penton base complex.

2. Modification of HSP interaction

The adenovirus fiber protein is a major determinant of adenovirus tropism (Gall *et al.* (1996) *J. Virol.* 70:2116-2123; Stevenson *et al.* (1995) *J. Virol.* 69:2850-2857). Dogma in the field has been that adenoviral entry occurs via binding to CAR and integrins. This is underscored by published data (Einfeld *et al.* (2001) *J. Virology* 75:11284-11291). It is shown herein, however, these published entry pathways are not the predominant ones that act *in vivo*. Moreover, as shown herein, the dominant entry pathway for hepatocytes *in vivo* involves a mechanism mediated by the fiber shaft, such as Ad5 shaft, through heparin sulfate proteoglycans binding.

It is shown herein that elimination of this binding eliminates entry via HSP binding, such as in hepatocytes. Adenoviral fiber shaft modifications that ablate viral interaction with HSP are provided. Thus, as provided herein, efficient detargeting of adenovirus *in vivo* can be achieved with appropriately designed fiber proteins. Suitable modifications, such as described herein, can be made with respect to any adenovirus in which the wild-type interacts with HSP.

As provided herein, the ability of an adenoviral vector to interact with HSP is modified. In particular, the ability to interact is reduced or eliminated. Modifications include insertions, deletions, individual amino acid mutations and

-29-

other mutations that alter the structure of the fiber shaft such that the HSP binding of the modified fiber protein is ablated when compared to the HSP binding of the wild-type fiber protein.

In a first aspect of this embodiment, an adenoviral fiber protein is
5 modified by mutating one or more of the amino acids that interact with HSP. For example, the HSP binding motif of the modified fiber protein is no longer able to interact with HSP on the cell surface, thus ablating the viral interaction with HSP. For example, the adenoviral fiber is from a subgroup C adenovirus. Binding to HSP can be eliminated or reduced by mutating the fiber shaft in order
10 to modify the ability of the HSP binding motif, which is, for example, KKTK sequence (SEQ ID No. 1) located between amino acid residues 91 to 94 in the Ad 5 fiber, to interact with HSP. The fiber proteins are modified by chemical and biological techniques known to those skilled in the art, such as site directed mutagenesis of nucleic acid encoding the fiber or other techniques as illustrated
15 herein.

In another aspect of this embodiment, the ability of a fiber to interact with HSP is modified by replacing the wild-type fiber shaft with a fiber shaft, or portion thereof, of an adenovirus that does not interact with HSP to produce chimeric fiber proteins. The portion is sufficient to reduce or eliminate
20 interaction with HSP. Examples of adenoviruses having fiber shafts that do not interact with HSP include (a) adenoviruses of subgroup B, such as, but are not limited to, Ad3, Ad35, Ad7, Ad11, Ad16, Ad21, Ad34, which do not have interaction with HSP, (b) adenoviruses of subgroup F, such as, but are not limited to, Ad40 and Ad41, specifically the short fiber, and (c) adenoviruses of
25 subgroup D, such as but are not limited to, Ad46. In another embodiment, adenoviral fiber shaft modifications that ablate viral interaction with HSP in combination with adenoviral fiber knob modifications that ablate viral interactions with CAR are provided. Suitable adenoviral fiber modifications include the fiber knob modifications are known to those of skill in the art and are exemplified
30 herein (see, also, US. Patent Application Serial No. 09/870,203, filed on 30 May 2001, and published as U.S. Published application No. 20020137213, in International Patent Application No. PCT/EP01/06286 filed on 1 June 2001).

-30-

Modifications of the fiber include mutations of at least one amino acid in the CD loop of a wild-type fiber protein of an adenovirus from subgroup C, D, or E, or the long wild-type fiber of an adenovirus from subgroup F, whereby the ability of a fiber protein to bind to CAR is reduced or substantially eliminated. The fiber proteins with ablated CAR interaction are modified by chemical and biological techniques known to those skilled in the art, as illustrated herein and as described in the above patent application.

Alternatively, adenoviral fiber modifications are made by replacing the wild-type fiber knob with a fiber knob of an adenovirus that does not interact with CAR. The fiber protein also will be selected so that it does not interact with HSP. Examples of adenoviruses having fiber knobs that do not interact with CAR include (a) adenoviruses of subgroup B, e.g., Ad3, Ad35, Ad7, Ad11, Ad16, Ad21, Ad34, (b) adenoviruses of subgroup F, e.g., Ad40 and Ad41, specifically the short fiber.

In another embodiment, adenoviral fiber shaft modifications that ablate viral interaction with HSP in combination with penton modifications that ablate viral interactions with α_v integrins are provided. Suitable adenoviral penton modifications include the penton modifications, which are well known to those of skill in the art (see, e.g., U.S. Patent No. 5,731,190; see, also Einfeld *et al.* (2001) *J. Virology* 75:11284-11291; and Bai *et al.* (1993) *J. Virology* 67:5198-5205).

For example, penton interaction with α_v integrins can be ablated (reduced or eliminated) by substitution of the RGD tripeptide motif, required for α_v interaction, in penton with a different tripeptide that does not interact with an α_v integrin. The penton proteins with ablated α_v integrin interactions are modified by chemical and biological techniques known to those skilled in the art (see, e.g., described U.S. Patent No. 6,731,190 and as illustrated herein). Generally, the adenovirus is a subgroup B or C adenovirus.

Also provided are adenoviral fiber shaft modifications that ablate viral interaction with HSP in combination with adenoviral fiber knob modifications that ablate viral interactions with CAR and with penton modifications that ablate viral interactions with α_v integrins. These modifications are described above and

-31-

prepared using chemical and biological techniques known to those skilled in the art and as illustrated herein. Generally the adenovirus is a subgroup B or subgroup C adenovirus.

Preparation of fibers modified to eliminate or reduce HSP interactions and
5 fibers modified to alter interactions with other receptors and cell surface proteins, such as CAR and/or α_v integrin, is also described in the Examples below. The nucleic acid and/or amino acid sequences of exemplary modified fibers, whose construction are described below) are set forth as SEQ ID Nos. 45-72 as follows:

10 SEQ ID Nos. 45 and 46 set forth the encoding nucleotide sequence and amino acid sequence of the modified fiber designated 5FKO1, where 5F refers to Adenovirus 5 fiber, KO1 is an exemplary mutation of the CAR interaction site described herein;

SEQ ID Nos. 47 and 48 set forth the encoding nucleotide sequence and
15 amino acid sequence of the modified ber designated 5FKO1RGD, which further includes an RGD ligand to demonstrate retargeting;

SEQ ID Nos. 49 and 50 set forth the encoding nucleotide sequence and amino acid sequence of the modified fiber designated 5FKO12, where 5F refers to Adenovirus 5 fiber, KO12 is another exemplary mutation of the CAR
20 interaction site described herein;

SEQ ID Nos. 51 and 52 set forth the encoding nucleotide sequence and amino acid sequence of the modified fiber designated 5F S* nuc, where 5F refers to Adenovirus 5 fiber, S* is an exemplary mutation of the shaft that alters binding to HSP;

25 SEQ ID Nos. 53 and 54 set forth the encoding nucleotide sequence and amino acid sequence of the modified fiber designated 5F S*RGD nuc, which further includes an RGD ligand;

SEQ ID Nos. 55 and 56 set forth the encoding nucleotide sequence and amino acid sequence of the modified ber designated 5FKO1S*, which contain
30 the KO1 and S* mutations;

-32-

SEQ ID Nos. 57 and 58 set forth the encoding nucleotide sequence and amino acid sequence of the modified fiber designated 5FKO1S*RGD, which further includes an RGD ligand;

5 SEQ ID Nos. 59 and 60 set forth the encoding nucleotide sequence and amino acid sequence of a Ad35 fiber;

SEQ ID Nos. 61 and 62 set forth the encoding nucleotide sequence and amino acid sequence of the modified fiber designated 35FRGD, which is 35F fiber with an RGD ligand;

10 SEQ ID Nos. 63 and 64 set forth the encoding nucleotide sequence and amino acid sequence of a Ad41 short fiber;

SEQ ID Nos. 65 and 66 set forth the encoding nucleotide sequence and amino acid sequence of the modified fiber designated 41sFRGD, which is 41F short fiber with an RGD ligand;

15 SEQ ID Nos. 67 and 68 set forth the encoding nucleotide sequence and amino acid sequence of Ad5 penton;

20 SEQ ID Nos. 69 and 70 set forth the encoding nucleotide sequence and amino acid sequence of the modified fiber designated 5TS35H, which is a chimeric fiber in which an Ad5 fiber tail and shaft regions (5TS; amino acids 1 to 403) are connected to an Ad35 fiber head region (35H; amino acids 137 to 323) to form the 5TS35H chimera; and

25 SEQ ID Nos. 71 and 72 set forth the encoding nucleotide sequence and amino acid sequence of the modified fiber designated 35TS5H, which is a chimeric fiber in which an Ad35 fiber tail and shaft regions (35TS; amino acids 1 to 136) are connected to an Ad5 fiber head region (5H; amino acids 404 to 581) to form the 35TS5H chimera.

30 SEQ ID No. 1 sets forth the nucleotide sequence of Ad fiber; SEQ ID Nos. 2 and 3 also set forth the coding nucleic acid sequences for fibers with modified fiber knobs for ablated CAR interaction (see, SEQ ID No. 2 for KO1 and SEQ ID No. 3 for KO12); SEQ ID No. 4 also sets for the encoding nucleic acid sequence of a modified penton for ablated α_v integrins (SEQ ID No. 4).

The modified fibers are displayed on virus particles by modifying the fiber protein and optionally additional proteins. This can be achieved by preparing

-33-

adenoviral vectors that express the modified capsid proteins and produce particles with modified fibers, or by packaging adenoviral vectors, particularly those that do not encode one or more capsid proteins in appropriate packaging lines. Hence, as discussed in detail below, adenoviral vectors and viral particles with modified fibers that do not bind to HSP are provided.

C. Nucleic acids, Adenoviral vectors and cells containing the nucleic acids and cells containing the vectors

Also provided are polynucleotides that encode modified capsid proteins and that encode vectors for preparation of adenovirus that express modified capsid proteins provided herein. The sequences of the wild-type adenovirus proteins are well known in the art and are modified as described herein. Nucleic acid molecules, such as cDNA that encode an exemplary modified fiber knob for ablated CAR interaction (see, SEQ ID No. 2 for KO1 and SEQ ID No. 3 for KO12) and for a modified penton for ablated α_v integrins (SEQ ID No. 4) are provided.

As discussed above, modified capsid proteins with altered tropism for CAR and α_v integrins are known and described in the patents, applications and literature cited herein and known to those of skill in the art (see, *e.g.*, U.S. Patent No. 5,731,190, U.S. application Serial No. 09/870,203, published as U.S. Published application No. 20020137213; and Bai *et al.* (1993) *J. Virology* 67:5198-5208).

Also provided are vectors including the polynucleotides provided herein. Such vectors include partial or complete adenoviral genomes and plasmids. Such vectors are constructed by techniques known to those skilled in the art and as illustrated herein. Also provided are adenoviral vectors modified by replacing whole fiber protein, or portions thereof, with the fiber proteins, or appropriate portions thereof, of an adenovirus that does not interact with HSP.

Adenoviruses that do not interact with HSP can be identified by using the methods described herein which detect binding or non-binding of fiber proteins and adenoviruses with HSP. Among the adenoviral vectors provided herein are those of subgroup C, which include Ad2 and Ad5, in which the nucleic acid encoding the fiber shaft or a portion including the HSP-binding portion is

-34-

replaced with nucleic acid encoding fiber or an appropriate portion thereof from a serotype, such as Ad35.

Adenoviral fiber modifications, thus, can be made in viral particles by replacing the entire fiber protein with the fiber protein of an adenovirus that does not interact with CAR and/or replacing the HSP binding portion with a portion that does not bind. Generally the adenovirus is a subgroup B or subgroup C adenovirus, and also an adenovirus of subgroup D, such as Ad46. Adenoviral vectors of subgroup C, such as Ad2 and Ad5, having a replaced fiber knob are prepared using techniques well known in the art and as illustrated herein.

1. Preparation of viral particles

The packaging cells used to produce the viruses provided herein contain the nucleic acid encoding the capsid protein, including the mutated fiber protein provided herein. Such nucleic acid can be transfected into the cell, generally part of as part of plasmid, or it can be infected into the cell with a viral vector. It can be stably incorporated into the genome of the cell, thus providing for a stable cell line. Alternatively, nucleic acid encoding the mutated capsid protein can be removed from the genome, in which case a transient complementing cell is employed.

The adenovirus genome to be packaged is transferred into the complementing cell by techniques known to those skilled in the art. These techniques include transfection or infection with the adenovirus. The nucleic acid encoding the mutated fiber protein can be in this genome instead of in the packaging cell.

In certain cases, when the nucleic acid encoding the mutated fiber is in the genome to be packaged, it can be desirable for the packaging cell to also encode a fiber protein. Such protein can assist in the maturation and packaging of an infectious particle. Such protein can be a wild-type fiber protein or one modified such that it is unable to attach to the penton base protein and is for use, for example, in producer cells where the fiber is included to provide the packaging function and the vector encodes a full-length fiber.

The packaging cells are cultured under conditions that permit the production of the desired viral particle. The viral particles are recovered by

-35-

standard techniques. An exemplary method for producing adenoviral particles provided herein is as follows. The nucleic acid encoding the mutated fiber protein is made using standard techniques in an adenoviral shuttle plasmid. This plasmid contains the right end of the virus, in particular from the end of the E3 region through the right ITR. This plasmid is co-transfected into competent cells of an *E. coli* strain, such as the well known *E. coli* strain BJ5183 (see, *e.g.*, Degryse (1996) *Gene* 170:45-50) along with a plasmid, which contains the remaining portion of the adenovirus genome, except for the E1 region and sometimes also the E2a region and also contains a corresponding region of homology. Homologous recombination between the two plasmids generates a full-length plasmid encoding the entire adenoviral vector genome.

This full-length adenoviral vector genome plasmid is then transfected into a complementing cell line. The transfection can be performed in the presence of a reagent that directs adenoviral particle entry into producer cells. Such reagents include, but are not limited to, polycations and bifunctional reagents, such as those described herein. A complementing cell is, for example, is a cell of the PER.C6 cell line, which contains the adenoviral E1 gene (PER.C6 is available, for example, from Crucell, The Netherlands; deposited under ECACC accession no. 96022940; see, also Fallaux *et al.* (1998) *Hum. Gene Ther.* 9:1909-1907; see, also, U.S. Patent No. 5,994,128) or an AE1-2a cell (see, Gorziglia *et al.* (1996) *J. Virology* 70:4173-4178; and and Von Seggern *et al.* (1998) *J. Gen. Virol.* 79:1461-1468)).

AE1-2a cells are derivatives of the A549 lung carcinoma line (ATCC # CCL 185) with chromosomal insertions of the plasmids pGRE5-2.E1 (also referred to as GRE5-E1-SV40-Hygro construct and listed in SEQ ID No. 41) and pMNeoE2a-3.1 (also referred to as MMTV-E2a-SV40-Neo construct and listed in SEQ ID No. 42), which provide complementation of the adenoviral E1 and E2a functions, respectively.

The 633 cell line (see, von Seggern *et al.* (2000) *J. Virology* 74:354-362), which stably expresses the adenovirus serotype 5 wild-type fiber protein, and was derived from the AE1-2a cell line, is another an example of complementing cells. When the cell line is 633 cells, the final passage of

-36-

adenoviral vector is performed on another complementing cell line (*e.g.*, Per.C6), which does not express wild-type Ad5 fiber.

The transfected complementing cells are maintained under standard cell culture conditions. The adenoviral plasmids recombine to form the adenoviral genome that is packaged. The particles are infectious, but replication deficient because their genome is missing at least the E1 genes. When performed in the 633 cells the particles contain wild-type and mutated fiber proteins. They are recovered from the crude viral lysate, amplified, and are purified by standard techniques.

The recovered particles can be used to infect PER.C6 or AE1-2a cells. This permits the recovery of particles whose capsids contain only the desired mutated fiber. This two-step procedure provides high titer batches of the adenoviral particles provided herein. The adenoviral particles can be replication competent or replication incompetent.

In one embodiment, the particles selectively replicate in certain predetermined target tissue but are replication incompetent in other cells and tissues. In a particular embodiment, the adenoviral particles replicate in abnormally proliferating tissue, such as solid tumors and other neoplasms. In replication conditional adenoviruses, a gene essential for replication is placed under control of a heterologous promoter which is cell or tissue specific. For example, the E1a gene is placed under control of a promoter which is active in a tumor cell to produce an oncolytic adenovirus or oncolytic adenoviral vector. Administration of oncolytic adenoviral vectors to tumor cells kills the tumor cells. Such replication conditional adenoviral particles and vectors can be produced by techniques known to those skilled in the art, such as those disclosed in the above-referenced U.S. Patent Nos. 5,998,205 and 5,801,029. These particles and vectors can be produced in adenoviral packaging cells as disclosed above. Generally packaging cells are those that have been designed to limit homologous recombination that could lead to wild-type adenoviral particles. Such cells are well known and include the packaging cell known as PER.C6 (see, *e.g.*, U.S. Patent Nos. 5,994,128 and 6,033,908; deposited under ECACC accession no. 96022940). Since oncolytic vectors are replication competent in certain cell

-37-

types, they can be amplified in cell lines derived from said cell type without provision of Ad complementary genes.

2. Adenoviral vectors and particles

- The adenovirus as used herein for production of the adenoviral vectors and particles can be of any serotype. Adenoviral stocks that can be employed as a source of adenovirus or adenoviral coat protein, such as fiber and/or penton base, can be amplified from the adenoviral serotypes 1 through 47, which are currently available from the American Type Culture Collection (ATCC, Rockville, Md.), or from any other serotype of adenovirus available from any other source.
- For instance, an adenovirus can be of subgroup A (e.g., serotypes 12, 18, 31), subgroup B (e.g., serotypes 3, 7, 11, 14, 16, 21, 34, 35), subgroup C (e.g., serotypes 1, 2, 5, 6), subgroup D (e.g., serotypes 8, 9, 10, 13, 15, 17, 19, 20, 22-30, 32, 33, 36-39, 42-47), subgroup E (serotype 4), subgroup F (serotype 40, 41), or any other adenoviral serotype.
- In certain embodiments, the adenovirus is a subgroup B or a subgroup C adenovirus. Subgroup C adenoviruses which are modified in as described herein, include, but are not limited to, Ad2 and Ad5. For Ad5, the mutation is made in the KKTK sequence (SEQ ID No. 1) located between amino acid residues 91 to 94. The fiber proteins can be modified by chemical and biological techniques known to those skilled in the art. These methods include, but are not limited to, site directed mutagenesis and techniques as illustrated herein.

- The adenoviral particle generally includes a targeting ligand as described above. The presence of the targeting ligand permits the delivery of a gene to a desired cell type which is different from the cell type that wild-type adenovirus particles infect or the same as that a wild-type particle infects, but allowing the infection in a selective manner, *i.e.*, non-target cell types are not significantly infected.

- The adenoviral vectors provided herein can be used to study cell transduction and gene expression *in vitro* or in various animal models. The latter case includes *ex vivo* techniques, in which cells are transduced *in vitro* and then administered to the animal. They also can be used to conduct gene therapy on humans or other animals. Such gene therapy can be *ex vivo* or *in vivo*. For *in*

-38-

vivo gene therapy, the adenoviral particles in a pharmaceutically-acceptable carrier are delivered to a human in a therapeutically effective amount in order to prevent, treat, or ameliorate a disease or other medical condition in the human through the introduction of a heterologous gene that encodes a therapeutic protein into cells in such human. The adenoviruses are delivered at a dose ranging from approximately 1 particle per kilogram of body weight to approximately 10^{14} particles per kilogram of body weight. Generally, they are delivered at a dose of approximately 10^6 particles per kilogram of body weight to approximately 10^{13} particles per kilogram of body weight, and typically the dose ranges from approximately 10^8 particles per kilogram of body weight to approximately 10^{12} particles per kilogram of body weight.

Any vectors known to those of skill in the art can be employed and used to produce viral particles that include fibers modified to ablate (including reduce) binding to HSP. Some exemplary vectors are as follows.

15 a. Gutless vectors

Gutted adenovirus vectors are those from which most or all viral genes have been deleted. They are grown by co-infection of the producing cells with a "helper" virus (such as using an E1-deleted Ad vector), where the packaging cells expresses the E1 gene products. The helper virus trans-complements the missing Ad functions, including production of the viral structural proteins needed for particle assembly. To incorporate the capsid modifications into a gutted adenoviral vector capsid, the changes must be made to the helper virus as described herein. All the necessary Ad proteins including the modified capsid protein are provided by the modified helper virus, and the gutted adenovirus particles are equipped with the particular modified capsid expressed by the host cells. The E1a, E1b, E2a, E2b and E4 are generally required for viral replication and packaging. If these genes are deleted, then the packaging cell must provide these genes or functional equivalents.

A helper adenovirus vector genome and a gutless adenoviral vector genome are delivered to packaging cells. The cells are maintained under standard cell maintenance or growth conditions, whereby the helper vector genome and the packaging cell together provide the complementing proteins for

-39-

the packaging of the adenoviral vector particle. Such gutless adenoviral vector particles are recovered by standard techniques. The helper vector genome can be delivered in the form of a plasmid or similar construct by standard transfection techniques, or it can be delivered through infection by a viral particle
5 containing the genome. Such viral particle is commonly called a helper virus. Similarly, the gutless adenoviral vector genome can be delivered to the cell by transfection or viral infection.

The helper virus genome can be the modified adenovirus vector genome as disclosed herein. Such genome also can be prepared or designed so that it
10 lacks the genes encoding the adenovirus E1A and E1B proteins. In addition, the genome can further lack the adenovirus genes encoding the adenovirus E3 proteins. Alternatively, the genes encoding such proteins can be present but mutated so that they do not encode functional E1A, E1B and E3 proteins. Furthermore, such vector genome can not encode other functional early proteins,
15 such as E2A, E2B3, and E4 proteins. Alternatively, the genes encoding such other early proteins can be present but mutated so that they do not encode functional proteins.

In producing the gutless vectors, the helper virus genome is also packaged, thereby producing helper virus. In order to minimize the amount of
20 helper virus produced and maximize the amount of gutless vector particles produced, the packaging sequence in the helper virus genome can be deleted or otherwise modified so that packaging of the helper virus genome is prevented or limited. Since the gutless vector genome will have an unmodified packaging sequence, it will be preferentially packaged.

25 One way to do this is to mutate the packaging sequence by deleting one or more of the nucleotides comprising the sequence or otherwise mutating the sequence to inactivate or hamper the packaging function. One exemplary approach is to engineer the helper genome so that recombinase target sites flank the packaging sequence and to provide a recombinase in the packaging cell. The
30 action of recombinase on such sites results in the removal of the packaging sequence from the helper virus genome. The recombinase can be provided by a nucleotide sequence in the packaging cell that encodes the recombinase. Such

-40-

sequence can be stably integrated into the genome of the packaging cell. Various kinds of recombinase are known by those skilled in the art, and include, but are not limited to, Cre recombinase, which operates on so-called lox sites, which are engineered on either side of the packaging sequence as discussed above (see, *e.g.*, U.S. Patent Nos. 5,919,676, 6,080,569 and 5,919,676; see, also, *e.g.*, Morsy and Caskey, *Molecular Medicine Today*, Jan. 1999, pgs. 18-24).

An example of a gutless vector is pAdARSVDys (Haecker *et al.* (1996) *Hum Gene Ther.* 7:1907-1914)). This plasmid contains a full-length human dystrophin cDNA driven by the RSV promoter and flanked by Ad inverted terminal repeats and packaging signals. 293 cells are infected with a first-generation Ad, which serves as a helper virus, and then transfected with purified pAdARSVDys DNA. The helper Ad genome and the pAdARSVDys DNA are replicated as Ad chromosomes, and packaged into particles using the viral proteins produced by the helper virus. Particles are isolated and the pAdARSVDys-containing particles separated from the helper by virtue of their smaller genome size and therefore different density on CsCl gradients. Other examples of gutless adenoviral vectors are known (see, *e.g.*, Sandig *et al.* (2000) *Proc. Natl. Acad. Sci. U.S.A.* 97(3):1002-7).

b. Oncolytic vectors

Briefly, oncolytic adenoviruses, which are viruses that replicate selectively in tumor cells, are designed to amplify the input virus dose due to viral replication in the tumor, leading to spread of the virus throughout the tumor mass. *In situ* replication of adenoviruses leads to cell lysis. This *in situ* replication permits relatively low, non-toxic doses to be highly effective in the selective elimination of tumor cells. One approach to achieving selectivity is to introduce loss-of-function mutations in viral genes that are essential for growth in non-target cells but not in tumor cells. (See, *e.g.*, U.S. Patent No. 5,801,029.) This strategy is exemplified by the use of Add1520, which has a deletion in the E1b-55KD gene. In normal cells, the adenoviral E1b-55KD protein is needed to bind to p53 to prevent apoptosis. In p53-deficient tumor cells, E1b-55K binding

-41-

to p53 is unnecessary. Thus, deletion of E1b-55KD should restrict vector replication to p53-deficient tumor cells.

Another approach is to use tumor-selective promoters to control the expression of early viral genes required for replication (see, *e.g.*, International PCT application Nos. WO 96/17053 and WO 99/25860). Thus, in this approach the adenoviruses selectively replicate and lyse tumor cells if the gene that is essential for replication is under the control of a promoter or other transcriptional regulatory element that is tumor-selective.

For example oncolytic adenoviral vectors that contain a cancer selective regulatory region operatively linked to an adenoviral gene essential for adenoviral replication are known (see, *e.g.*, U.S. Patent No. 5,998,205). Adenoviral genes essential for replication include, but are not limited to, E1a, E1b, E2a, E2b and E4. For example, an exemplary oncolytic adenoviral vector has a cancer selective regulatory region operatively linked to the E1a gene. In other embodiments, the oncolytic adenoviral vector has a cancer selective regulatory region of the present invention operatively linked to the E1a gene and a second cancer selective regulatory region operatively linked to the E4 gene. The vectors also can include at least one therapeutic transgene, such as, but not limited to, a polynucleotide encoding a cytokine such as GM-CSF that can stimulate a systemic immune response against tumor cells.

Other exemplary oncolytic adenoviral vectors include those in which expression of an adenoviral gene, which is essential for replication, is controlled by E2F-responsive promoters, which are selectively transactivated in cancer cells. Thus, vectors that contains an adenoviral nucleic acid backbone that contains in sequential order: A left ITR, an adenoviral packaging signal, a termination signal sequence, an E2F responsive promoter which is operably linked to a first gene, such as E1a, essential for replication of the recombinant viral vector and a right ITR (see, published International PCT application No. WO02/06786, and U.S. Patent No. 5,998,205).

In other embodiments, the oncolytic adenoviral vector has a cancer selective regulatory region operatively linked to the E1a gene and a second cancer selective regulatory region operatively linked to the E4 gene. The vectors

-42-

can also carry at least one therapeutic transgene, such as, but not limited to, a polynucleotide encoding a cytokine such as GM-CSF that can stimulate a systemic immune response against tumor cells.

3. Packaging

5 The viral particles provided herein can be made by any method known to those of skill in the art. Generally they are prepared by growing the adenovirus vector that contains nucleic acid that encodes the modified fiber protein in a standard adenovirus packaging cells to produce particles that express the modified fibers. Alternatively, the vectors do not encode fibers. Such vectors
10 are packaged in producer cells to produce particles that express the modified fiber proteins.

As discussed, recombinant adenoviral vectors generally have at least a deletion in the first viral early gene region, referred to as E1, which includes the E1a and E1b regions. Deletion of the viral E1 region renders the recombinant
15 adenovirus defective for replication and incapable of producing infectious viral particles in subsequently-infected target cells. Thus, to generate E1-deleted adenovirus genome replication and to produce virus particles requires a system of complementation which provides the missing E1 gene product. E1
20 complementation is typically provided by a cell line expressing E1, such as the human embryonic kidney packaging cell line, i.e. an epithelial cell line, called 293. Cell line 293 contains the E1 region of adenovirus, which provides E1 gene region products to "support" the growth of E1-deleted virus in the cell line (see, e.g., Graham *et al.*, *J. Gen. Virol.* 36: 59-71, 1977). Additionally, cell lines
25 that may be usable for production of defective adenovirus having a portion of the adenovirus E4 region have been reported (WO 96/22378). Multiply deficient adenoviral vectors and complementing cell lines have also been described (WO 95/34671, U.S. Patent No. 5,994,106).

For example, copending U.S. application Serial No. 09/482,682 (also filed as International PCT application No. PCT/EP00/00265, filed January 14,
30 200, published as International PCT application No. WO/0042208) provides packaging cell lines that support viral vectors with deletions of major portions of the viral genome, without the need for helper viruses and also provides cell lines

-43-

and helper viruses for use with helper-dependent vectors. The packaging cell line has heterologous DNA stably integrated into the chromosomes of the cellular genome. The heterologous DNA sequence encodes one or more adenovirus regulatory and/or structural polypeptides that complement the genes deleted or mutated in the adenovirus vector genome to be replicated and packaged.

Packaging cell lines express, for example, one or more adenovirus structural proteins, polypeptides, or fragments thereof, such as penton base, hexon, fiber, polypeptide IIIa, polypeptide V, polypeptide VI, polypeptide VII, polypeptide VIII, and biologically active fragments thereof. The expression can be constitutive or under the control of a regulatable promoter. These cell lines are particularly designed for expression of recombinant adenoviruses intended for delivery of therapeutic products. For use herein, such packaging cell lines can express the modified capsid proteins, such as the fiber proteins whose binding to HSP is reduced or eliminated, and/or the modified penton and hexon proteins.

Particular packaging cell lines complement viral vectors having a deletion or mutation of a DNA sequence encoding an adenovirus structural protein, regulatory polypeptides E1A and E1B, and/or one or more of the following regulatory proteins or polypeptides: E2A, E2B, E3, E4, L4, or fragments thereof.

The packaging cell lines are produced by introducing each DNA molecule into the cells and then into the genome via a separate complementing plasmid or plurality of DNA molecules encoding the complementing proteins can be introduced via a single complementing plasmid. Of interest herein, is a variation in which the complementing plasmid includes DNA encoding adenovirus fiber protein (or a chimeric or modified variant thereof), from Ad virus of subgroup D, such as Ad 37, polypeptide or fragment thereof.

For applications, such as therapeutic applications, the delivery plasmid further can include a nucleotide sequence encoding a heterologous polypeptide. Exemplary delivery plasmids include, but are not limited to, pDV44, pΔE1Bβ-gal and pΔE1sp1B. In a similar or analogous manner, therapeutic nucleic acids, such as nucleic acids that encode therapeutic genes, can be introduced.

-44-

The cell further includes a complementing plasmid encoding a fiber as contemplated herein; the plasmid or portion thereof is integrated into a chromosome(s) of the cellular genome of the cell.

Typically, the packaging cell lines will contain nucleic acid encoding the
5 fiber protein or modified protein stably integrated into a chromosome or
chromosomes in the cellular genome. The packaging cell line can be derived from
a procaryotic cell line or from a eukaryotic cell line. While various embodiments
suggest the use of mammalian cells, and more particularly, epithelial cell lines, a
variety of other, non-epithelial cell lines are used in various embodiments. Thus,
10 while various embodiments disclose the use of a cell line selected from among
the 293, A549, W162, HeLa, Vero, 211, and 211A cell lines, and any other cell
lines suitable for such use are likewise contemplated herein.

D. Addition of a targeting ligand

The viral particles that are detargeted as described herein, can be
15 retargeted to selected cells and/or tissues by inclusion of an appropriate
targeting ligand in the capsid. The ligand can be included in any of the capsid
proteins, such as fiber, hexon and penton. Loci for inclusion of nucleic acid
encoding a is known to those of skill in the art for a variety of adenovirus
serotypes; if necessary appropriate loci and other parameters can be empirically
20 determined.

The ligand can be produced as a fusion by inclusion of the coding
sequences in the nucleic acid encoding a capsid protein, or chemically
conjugated, such as via ionic, covalent or other interactions, to the capsid or
bound to the capsid (*e.g.*, by Ab-ligand fusion, where Ab binds capsid protein; or
25 by disulfide bonding or other crosslinking moieties or chemistries).

Thus, for example, a modified fiber nucleic acid also can include
sequences of nucleotides that encode a targeting ligand to produce viral particles
that include a targeting ligand in the capsid. Targeting ligand and methods for
including such ligands in viral capsids are well known. For example, inclusion of
30 targeting ligands in fiber proteins is described in U.S. Patent Nos. 5,543,328 and
5,756,086 and in U.S. Patent Application Serial No. 09/870,203, published as
U.S. Published application No. 20020137213, and International Patent

-45-

Application No. PCT/EP01/06286. For different serotypes and strains of adenoviruses, loci for insertion of targeting ligands can be empirically determined. For different serotypes and strains, such loci can vary.

Because the adenovirus fiber has a trimeric structure, the ligand can be
5 selected or designed to have a trimeric structure so that up to three molecules of the ligand are present for each mature fiber. Such ligands can be incorporated into the fiber protein using methods known in the art (see, *e.g.*, U.S. Patent No. 5,756,086). Instead of the fiber, the targeting ligand can be included in the penton or hexon proteins. Inclusion of targeting ligands in penton (see for
10 example, in U.S. Patent Nos. 5,731,190 and 5,965,431) and in hexon (see for example, in U.S. Patent No. 5,965,541) is known.

In one exemplary embodiment, the ligand is included in a fiber protein, which is a fiber protein mutated as described herein. As shown herein, the targeting ligand can be included, for example, within the HI loop of the fiber
15 protein. Any ligand that can fit in the HI loop and still provide a functional virus is contemplated herein. Such ligands can be as long as or longer than 80-100 amino acids (see, *e.g.*, Belousova *et al.* (2002) *J. Virol.* 76:8621-8631). Such ligands are added by techniques known in the art (see, *e.g.*, published International Patent Application publication No. WO99/39734 and U.S. Patent
20 Application number 09/482,682). Other ligands can be discovered through techniques known to those skilled in the art. Some non-limiting examples of these techniques include phage display libraries or by screening other types of libraries.

Targeting ligands include any chemical moiety that preferentially directs
25 an adenoviral particle to a desired cell type and/or tissue. The categories of such ligands include, but are not limited to, peptides, polypeptides, single chain antibodies, and multimeric proteins. Specific ligands include the TNF superfamily of ligands which include tumor necrosis factors (or TNF's) such as, for example, TNF α and TNF β , lymphotoxins (LT), such as LT- α and LT- β , Fas ligand which
30 binds to Fas antigen; CD40 ligand, which binds to the CD40 receptor of B-lymphocytes; CD30 ligand, which binds to the CD30 receptor of neoplastic cells of Hodgkin's lymphoma; CD27 ligand, NGF ligand, and OX-40 ligand;

-46-

transferrin, which binds to the transferrin receptor located on tumor cells, activated T -cells, and neural tissue cells; ApoB, which binds to the LDL receptor of liver cells; alpha-2-macroglobulin, which binds to the LRP receptor of liver cells; alpha-1 acid glycoprotein, which binds to the asialoglycoprotein receptor of liver; mannose-containing peptides, which bind to the mannose receptor of macrophages; sialyl-Lewis-X antigen-containing peptides, which bind to the ELAM-I receptor of activated endothelial cells; CD34 ligand, which binds to the CD34 receptor of hematopoietic progenitor cells; ICAM-I, which binds to the LFA-I (CD11b/CD18) receptor of lymphocytes, or to the Mac-I (CD11a/CD18) receptor of macrophages; M-CSF, which binds to the c-fms receptor of spleen and bone marrow macrophages; circumsporozoite protein, which binds to hepatic Plasmodium falciparum receptor of liver cells; VLA-4, which binds to the VCAM-I receptor of activated endothelial cells; HIV gp120 and Class II MHC antigen, which bind to the CD4 receptor of T -helper cells; the LDL receptor binding region of the apolipoprotein E (ApoE) molecule; colony stimulating factor, or CSF, which binds to the CSF receptor; insulin-like growth factors, such as IGF-I and IGF-II, which bind to the IGF-I and IGF-II receptors, respectively; Interleukins 1 through 14, which bind to the Interleukin 1 through 14 receptors, respectively; the Fv antigen-binding domain of an immunoglobulin; gelatinase (MMP) inhibitor; bombesin, gastrin-releasing peptide; substance P; somatostatin; luteinizing hormone releasing hormone (LHRH); vasoactive peptide (VIP); gastrin; melanocyte stimulating hormone (MSH); cyclic RGD peptide and any other ligand or cell surface protein-binding (or targeting) molecule.

E. Heterologous polynucleotides and Therapeutic Nucleic Acids

The packaged adenoviral genome also can contain a heterologous polynucleotide that encodes a product of interest, such as a therapeutic protein. Adenoviral genomes containing heterologous polynucleotides are well known (see, e.g., U.S. Patent Nos. 5,998,205, 6,156,497, 5,935,935, and 5,801,029). These can be used for *in vitro* and *in vivo* delivery of the products of heterologous polynucleotides or the heterologous polynucleotides.

Thus, the adenoviral particles provided herein can be used to engineer a cell to express a protein that it otherwise does not express or does not express

-47-

in sufficient quantities. This genetic engineering is accomplished by infecting the desired cell with an adenoviral particle whose genome includes a desired heterologous polynucleotide. The heterologous polynucleotide is then expressed in the genetically engineered cells. For use herein the cell is generally a

5 mammalian cell, and is typically a primate cell, including a human cell. The cell can be inside the body of the animal (*in vivo*) or outside the body (*in vitro*). Heterologous polynucleotides (also referred to as heterologous nucleic acid sequences) are included in the adenoviral genome within the particle and are added to that genome by techniques known in the art. Any heterologous

10 polynucleotide of interest can be added, such as those disclosed in U.S. Patent No. 5,998,205, incorporated herein by reference. Polynucleotides that are introduced into an Ad genome or vector can be any that encode a protein of interest or that are regulatory sequences. Proteins include, but are not limited to, therapeutic proteins, such as an immunostimulating protein, such as an

15 interleukin, interferon, or colony stimulating factor, such as granulocyte macrophage colony stimulating factor (GM-CSF; see, *e.g.*, 5,908,763F. Generally, such GM-CSF is a primate GM-CSF, including human GM-CSF. Other immunostimulatory genes include, but are not limited to, genes that encode cytokines IL1, IL2, IL4, IL5, IFN, IFN, TNF, IL12, IL18, and flt3), proteins that

20 stimulate interactions with immune cells (B7, CD28, MHC class I, MHC class II, TAPs), tumor-associated antigens (immunogenic sequences from MART-1, gp100(pmel-17), tyrosinase, tyrosinase-related protein 1, tyrosinase-related protein 2, melanocyte-stimulating hormone receptor, MAGE1, MAGE2, MAGE3, MAGE12, BAGE, GAGE, NY-ESO-1, -catenin, MUM-1, CDK-4, caspase 8, KIA

25 O205, HLA-A2R1701, -fetoprotein, telomerase catalytic protein, G-250, MUC-1, carcinoembryonic protein, p53, Her2/neu, triosephosphate isomerase, CDC-27, LDLR-FUT, telomerase reverse transcriptase, and PSMA), cDNAs of antibodies that block inhibitory signals (CTLA4 blockade), chemokines (MIP1, MIP3, CCR7 ligand, and calreticulin), and other proteins.

30 Other polynucleotides, including therapeutic nucleic acids, such as therapeutic genes, of interest include, but are not limited to, anti-angiogenic, and suicide genes. Anti-angiogenic genes include, but are not limited to, genes that

-48-

- encode METH-1, METH -2, TrpRS fragments, proliferin-related protein, prolactin fragment, PEDF, vasostatin, various fragments of extracellular matrix proteins and growth factor/cytokine inhibitors. Various fragments of extracellular matrix proteins include, but are not limited to, angiostatin, endostatin, kininostatin, fibrinogen-E fragment, thrombospondin, tumstatin, canstatin, and restin. Growth factor/cytokine inhibitors include, but are not limited to, VEGF/VEGFR antagonist, sFlt-1, sFlk, sNRP1, angiopoietin/tie antagonist, sTie-2, chemokines (IP-10, PF-4, Gro-beta, IFN-gamma (Mig), IFN, FGF/FGFR antagonist (sFGFR), Ephrin/Eph antagonist (sEphB4 and sephrinB2), PDGF, TGF and IGF-1.
- 10 A "suicide gene" encodes a protein that can lead to cell death, as with expression of diphtheria toxin A, or the expression of the protein can render cells selectively sensitive to certain drugs, e.g., expression of the Herpes simplex thymidine kinase gene (HSV-TK) renders cells sensitive to antiviral compounds, such as acyclovir, gancyclovir and FIAU (1-(2-deoxy-2-fluoro--beta.-
- 15 D-arabinofuranosyl)-5-iodouracil). Other suicide genes include, but are not limited to, genes that encode carboxypeptidase G2 (CPG2), carboxylesterase (CA), cytosine deaminase (CD), cytochrome P450 (cyt-450), deoxycytidine kinase (dCK), nitroreductase (NR), purine nucleoside phosphorylase (PNP), thymidine phosphorylase (TP), varicella zoster virus thymidine kinase (VZV-TK), and
- 20 xanthine-guanine phosphoribosyl transferase (XGPRT). Alternatively, a therapeutic nucleic acid can exert its effect at the level of RNA, for instance, by encoding an antisense message or ribozyme, a protein that affects splicing or 3' processing (e.g., polyadenylation), or a protein that affects the level of expression of another gene within the cell, e.g. by mediating an altered rate of
- 25 mRNA accumulation, an alteration of mRNA transport, and/or a change in post-transcriptional regulation. The addition of a therapeutic nucleic acid to a virus results in a virus with an additional antitumor mechanism of action. Thus, a single entity (i.e., the virus carrying a therapeutic transgene) is capable of inducing multiple antitumor mechanisms. Other encoded proteins, include, but
- 30 are not limited to, herpes simplex virus thymidine kinase (HSV-TK), which is useful as a safety switch (see, U.S. Patent Application No. 08/974,391, filed

November 19, 1997, which published as PCT Publication No. WO/9925860), Nos, FasL, and sFasR (soluble Fas receptor).

Also contemplated are combinations of two or more transgenes with synergistic, complementary and/or nonoverlapping toxicities and methods of action. The resulting adenovirus can retain the viral oncolytic functions and, for example, additionally are endowed with the ability to induce immune and anti-angiogenic responses and other responses as desired.

Therapeutic polynucleotides and heterologous polynucleotides also include those that exert an effect at the level of RNA or protein. These include include a factor capable of initiating apoptosis, RNA, such as RNAi and other double-stranded RNA, antisense and ribozymes, which among other capabilities can be directed to mRNAs encoding proteins essential for proliferation, such as structural proteins, transcription factors, polymerases, genes encoding cytotoxic proteins, genes that encode an engineered cytoplasmic variant of a nuclease (e.g. RNase A) or protease (e.g. trypsin, papain, proteinase K and carboxypeptidase). Other polynucleotides include a cell or tissue specific promoters, such as those used in oncolytic adenoviruses (see, *e.g.*, U.S. Patent No. 5,998,205).

The heterologous polynucleotide encoding a polypeptide also can contain a promoter operably linked to the coding region. Generally the promoter is a regulated promoter and transcription factor expression system, such as the published tetracycline-regulated systems, or other regulatable systems (WO 01/30843), to allow regulated expression of the encoded polypeptide.

Exemplary of other promoters, are tissue-selective promoters, such as those described in U.S. Patent No. 5,998,205. An exemplary regulatable promoter system is the Tet-On(and Tet-Off(systems currently available from Clontech (Palo Alto, CA). This promoter system allows the regulated expression of the transgene controlled by tetracycline or tetracycline derivatives, such as doxycycline. This system can be used to control the expression of the encoded polypeptide in the viral particles and nucleic acids provided herein. Other regulatable promoter systems are known (see, *e.g.*, published U.S. No. 20020168714, entitled "Regulation of Gene Expression Using Single-Chain,

-50-

Monomeric, Ligand Dependent Polypeptide Switches," which describes gene switches that contain ligand binding domains and transcriptional regulating domains, such as those from hormone receptors). Other suitable promoters that can be employed include, but are not limited to, adenoviral promoters, such as the adenoviral major late promoter and/or the E3 promoter; or heterologous promoters, such as the cytomegalovirus (CMV) promoter; the Rous Sarcoma Virus (RSV) promoter; inducible promoters, such as the MMT promoter, the metallothionein promoter; heat shock promoters; the albumin promoter; and the ApoA1 promoter.

10 Therapeutic transgenes can be included in the viral constructs and resulting particles. Among these are those that result in an "armed" virus. For example, rather than delete E3 region as in some embodiments described herein, all or a part of the E3 region can be preserved or re-inserted in an oncolytic adenoviral vector (discussed above). The presence of all or a part of the E3
15 region can decrease the immunogenicity of the adenoviral vector. It also increases cytopathic effect in tumor cells and decreases toxicity to normal cells. Typically such vector expresses more than half of the E3 proteins.

 Adenoviruses for therapy, including those for human therapy, are known. Such known viruses can be modified as provided herein to reduce or eliminate
20 interaction with HSPs and optionally additional receptors. The adenoviral vectors that are used to produce the viral particles can include other modifications. Modifications include modifications to the adenovirus genome that is packaged in the particle in order to make an adenoviral vector. As discussed above, adenovirus vectors and particles with a variety of modifications are available.
25 Modifications to adenoviral vectors include deletions known in the art, such as deletions in one or more of the E1, E2a, E2b, E3, or E4 coding regions. These adenoviruses are sometimes referred to as early generation adenoviruses. include those with deletions of all of the coding regions of the adenoviral genome ("gutless" adenoviruses, discussed above) and also include replication-conditional adenoviruses, which are viruses that replicate in certain types of cells or
30 tissues but not in other types as a result of placing adenoviral genes essential for replication under control of a heterologous promoter (discussed above; see, also

-51-

U.S. Patent No. 5,998,205, U.S. Patent No. 5,801,029; U.S. patent application 60/348,670 and corresponding published International PCT application No. WO02/06786). These include the cytolytic, cytopathic viruses (or vectors), including the oncolytic viruses discussed above.

5 Alternatively, as discussed above, the vector can include a mutation or deletion in the E1b gene. Typically such mutation or deletion in the E1b gene is such that the E1b-19kD protein becomes non-functional. This modification of the E1b region can be combined with vectors where all or a part of the E3 region is present.

10 The oncolytic adenoviral vector can further include at least one heterologous coding sequence, such as one that encodes a therapeutic product. The heterologous coding sequence, such as therapeutic gene, is generally, although not necessarily, in the form of cDNA, and can be inserted at any locus that does not adversely affect the infectivity or replication of the vector. For
15 example, it can be inserted in an E3 region in place of at least one of the polynucleotide sequences that encode an E3 protein, such as, for example, the 19kD or 14.7 kD E3 gene.

F. Propagation and Scale-up

20 Since doubly ablated adenoviral vectors containing mutations in the fiber and/or penton capsid proteins result in inefficient cell binding and entry via the CAR/ α v integrin entry pathway, scaled up technologies improve the growth and propagation of such vectors to produce high titers of the adenoviral vectors for clinical use. Thus, also provided is a method for scaling up the production of detargeted adenoviral vectors. The detargeted adenoviral vectors comprise an
25 adenoviral vector modified to ablate the interaction of said vector with at least one host cell receptor compared with a wild-type adenoviral vector. The detargeted adenoviral vectors can comprise an adenoviral vector modified to ablate the interaction of said vector with one, two, three or more host cell receptors. Thus, the method is suitable for producing the detargeted adenoviral
30 vectors disclosed herein.

As noted, growth and propagation of doubly and fully ablated adenoviral vectors is enhanced by new scale up technologies. Doubly ablated vectors

-52-

contain mutations in the fiber and penton capsid proteins that result in inefficient cell binding and entry via the normal cellular entry pathway using CAR and integrins. These vectors are fully detargeted *in vitro* and, thus, alternative cellular entry strategies allow for the efficient growth and generation of high titer preparations.

Two strategies have been envisioned to scale up vectors that are detargeted via fiber and/or penton modifications. These include: (a) the use of pseudoreceptor cell lines engineered to express a surface receptor that binds a ligand displayed on the vector (see, *e.g.*, International PCT application No. WO 98/54346) and (b) complementing cell lines that are engineered to express native fiber and that can be engineered to express native fiber and penton (see, *e.g.*, International PCT application No. WO 00/42208). Although these systems have shown promise for scaling up ablated adenoviral vectors, there is a need to develop a system for the simple, efficient production of the fully detargeted adenoviral vector for therapeutic uses.

Provided herein is a scale-up method for the propagation of detargeted adenoviral vectors. The method uses polycations and/or bifunctional reagents, which when added to tissue culture medium, bind adenoviral particles and direct their entry into the producer cells.

Reagents (also called medium additives) also can be included in the tissue culture medium containing producer cells to be infected with the detargeted adenoviral vectors. Alternatively the reagents can be pre-mixed with the virus, which mixture is then added to the tissue producer cells. The reagents can be added to tissue culture medium containing producer cells, or producer cells can be added to tissue culture medium containing the reagents. Any suitable producer cell known to the skilled artisan can be used in the present methods. The reagents can be added at the same time that the producer cells are infected with detargeted adenoviral vectors. Generally the reagents are present in the tissue culture medium prior to infection by the detargeted adenoviral vectors. The medium additives are maintained in the tissue culture medium during vector growth, spread and propagation. High titer yields of adenoviral vectors are obtained by this method.

-53-

Reagents which are useful in this method are those that are capable of directing adenoviral particle entry into the producer cells. Such reagents include, but are not limited to, polycations and bifunctional reagents. Suitable polycations include, but are not limited to, polythetylenimine; protamine sulfate; 5 poly-L-lysine hydrobromide; poly(dimethyl diallyl ammonium) chloride (Merquat(r)-100, Merquat(r)280, Merquat(r)550); poly-L-arginine hydrochloride; poly-L-histidine; poly(4-vinylpyridine), poly(4-vinylpyridine) hydrochloride; poly(4-vinylpyridine)cross-linked, methylchloride quaternary salt; poly(4-vinylpyridine-co-styrene); poly(4-vinylpyridinium poly(hydrogen fluoride)); poly(4-vinylpyridinium-P-toluene sulfonate); poly(4-vinylpyridinium-tribromide); poly(4-vinylpyrrolidone-co-2-dimethylamino-ethyl methacrylate); polyvinylpyrrolidone, 10 cross-linked; poly vinylpyrrolidone, poly(melamine-co-formaldehyde); partially methylated; hexadimethrine bromide; poly(Glu, Lys) 1:4 hydrobromide; poly(Lys, Ala) 3:1 hydrobromide; poly(Lys, Ala) 2:1 hydrobromide; poly-L-lysine succinylated; poly(Lys, Ala) 1:1 hydrobromide; and poly(Lys, Trp) 1:4 15 hydrobromide.

Suitable bifunctional reagents include, but are not limited to, antibodies or peptides that bind to the adenoviral capsid and that also contain a ligand that allows interaction with specific cell surface receptors of the producer cells. 20 Examples of bifunctional reagents include: (a) anti-fiber antibody ligand fusions, (b) anti-fiber-Fab-FGF conjugate, (c) anti-penton-antibody ligand fusions, (d) anti-hexon antibody ligand fusions and (e) polylysine-peptide fusions. The ligand is any ligand that will bind to any cell surface receptor found on the producer cells.

25

The following examples are included for illustrative purposes only and are not intended to limit the scope of the invention.

-54-

EXAMPLE 1**Construction of Ad5 Vectors Containing the Fiber AB Loop, KO1 and Penton, PD1 Mutations and Derivatives Thereof**

Three recombinant adenoviral vectors were prepared that contain the KO1 fiber or PD1 penton base mutations either alone or in combination, these vectors are designated Av3nBgFKO1, Av1nBgPD1, and Av1nBgFKO1PD1. Construction of these vectors is described below and a general description of each vector is set forth in Table 1.

TABLE 1
Description Of Detargeted
Recombinant Adenoviral Vectors Used For Scale-up
Vector

Vector	Description
Av3nBg	An E1, E2a, E3-deleted adenoviral vector encoding a nuclear localizing β -galactosidase
Av1nBg	An E1 and E3-deleted adenoviral vector encoding a nuclear localizing β -galactosidase
Av3nBgFKO1	The same as Av3nBg but containing the KO1 mutation in the fiber gene
Av1nBgPD1	The same as Av1nBg but containing the PD1 mutation in the penton gene
Av1nBgFKO1PD1	The same as Av1nBg but containing the fiber KO1 and penton PD1 mutations

20 Av1nBg

This is a well-known vector, its sequence is set forth in SEQ ID No. 43.

Av3nBg

This is a well-known vector, its sequence is set forth in SEQ ID No. 44.

Av3nBgFKO1**25 Genetic incorporation of the KO1 fiber mutation to generate Av3nBgFKO1**

The adenoviral vector Av3nBgFKO1 was generated in an E1-, E2a-, E3-deleted backbone based on the adenovirus serotype 5 genome. It contains a RSV promoted nuclear-localizing β -galactosidase gene in place of the E1 region. In addition, the fiber gene carries the KO1 mutation. This mutation results in a

-55-

substitution of fiber amino acids 408 and 409, changing them from serine and proline to glutamic acid and alanine, respectively.

The vector was constructed as follows. First, the plasmid pSKO1 (Figure 1) was digested with the restriction enzymes SphI and MnlI. The resulting DNA fragments were separated by electrophoresis on an agarose gel. The 1601 bp fragment containing all but the 5' end of the fiber gene was excised from the agarose gel and the DNA was isolated and purified. The fragment was then ligated with the 9236 bp fragment of p5FloxHRFRGD, which had been digested with SphI and MnlI. The resulting plasmid, p5FloxHRFKO1, was digested with SpeI and PaeI and the 6867 bp fragment containing the fiber gene was isolated. The fragment was ligated with the 24,630 bp SpeI-PaeI fragment of pNDSQ3.1. The resulting plasmid, pNDSQ3.1KO1 (Figure 2), was used together with pAdmireRSVnBg (Figure 3A) to generate a plasmid which encodes the full-length adenoviral vector genome. It, however, was necessary to remove the PaeI site from pNDSQ3.1KO1 (Figure 2) prior to recombination with pAdmireRSVnBg (Figure 3A) so that the final plasmid contains a unique PaeI site adjacent to the 5' ITR. The PaeI site in pNDSQ3.1KO1 was removed by digestion with PaeI followed by blunting with T4 DNA Polymerase and religation. The resulting plasmid was called pNDSQ3.1KO1(Pac).

To generate a full-length plasmid containing the entire adenoviral genome, pAdmireRSVnBg (Figure 3A) was digested with Sall and co-transfected into competent cells of the *E. coli* strain BJ5183 along with pNDSQ3.1KO1ΔPac, which had been digested with BstBI. Homologous recombination between the two plasmids generated a full-length plasmid encoding the entire adenoviral vector genome, which was called pFLAv3nBgFKO1.

The plasmid pFLAv3nBgKO1 was linearized with PaeI and transfected into 633 cells. In the fiber complementing 633 cell line, the resulting viral DNA containing the KO1 mutation is capable of being packaged into infectious viral particles containing a mixture of wildtype fiber and mutant fiber proteins. After five rounds of amplification in 633 cells, a cytopathic effect was observed. Three more rounds of amplification in 633 cells were performed followed by purification of the virus by standard CsCl centrifugation procedures. This viral

-56-

preparation was used to infect AE1-2a cells, which do not express fiber. The resulting virus contained only the mutant fiber protein on its capsid. Virus particles were purified by standard CsCl centrifugation procedures.

Av1nBgFKO1

- 5 The vector Av1nBgFKO1 is made in a similar manner to Av3nBgFKO1 described above.

Av1nBgKO12

- An additional fiber AB loop mutation (described by Einfeld *et al.* (2001) *J. Virology* 75:11284-11291) was incorporated into the genome of Av1nBg. This
- 10 AB loop mutation is a four amino acid substitution, R512S, A515G, E516G, and K517G, and is referred to as KO12. The KO12 mutation was incorporated into the fiber gene by PCR gene overlap extension using the plasmid pSQ1 (Figure 3B) as template. The pSQ1 plasmid contains most of the Ad5 genome, extending from base pair 3329 through the right ITR, in a pBR322 backbone.
- 15 First, a segment of the Ad5 genome extending from within the E3 region into the fiber gene was amplified by PCR using the plasmid pSQ1 as a template with the following primers termed 5FF, 5'-GAA CAG GAG GTG AGC TTA GA-3' SEQ ID No. 4), and 5FR, 5'-TCC GCC TCC ATT TAG TGA ACA GTT AGG AGA TGG AGC TGG TGT G-3' (SEQ ID No. 6). The primer 5FR contains an 18 base
- 20 5'-extension that encodes the modified fiber AB loop amino acids from 512 through 517. A second PCR using pSQ1 as a template amplified the region immediately 3' of the AB loop substitution and extending past the MunI site located 40 base pairs 3' of the fiber gene stop codon. The two primers used for this reaction were 3FF: 5'-TCA CTA AAT GGA GGC GGA GAT GCT AAA CTC.
- 25 ACT TTG GTC TTA AC-3' (SEQ ID No. 7), and 3FR: 5'-GTG GCA GGT TGA ATA CTA GG-3' (SEQ ID No.8). The primer 3FR contains an 18 base 5'-extension that encodes the modified fiber AB loop amino acids 512 through 517. Amplified products of the expected size were obtained and used in a second PCR with the end primers 5FF and 3FR to join the fragments together.
- 30 The KO12 PCR fragment was digested with XbaI and MunI cloned directly into the fiber shuttle plasmid, pFBshuttle(EcoRI) to generate the plasmid pFBSEKO12 which contains the 8.8kB EcoRI fragment of pSQ1. The pFBSEKO12 plasmid

-57-

was digested with XbaI and EcoRI and cloned into pSQ1 using a three-way ligation to generate pSQ1KO12 (Figure 3C). The KO12 cDNA was incorporated into the genome of Av1nBg, an adenovirus vector with E1 and E3 deleted encoding β -galactosidase, by homologous recombination between ClaI-linearized

5 pSQ1KO12 and pAdmireRSVnBg digested with Sall and PacI to generate Av1nBgKO12. The KO12 vector was transfected in 633 cells, scaled-up on non-fiber expressing cells and purified, as described above for KO1.

Av1nBgPD1

Genetic incorporation of the PD1 penton mutation to generate Av1nBgPD1

10 The adenoviral vector Av1nBgPD1 is an E1-, E3-deleted vector based on the adenovirus serotype 5 genome. It contains a RSV promoted nuclear-localizing β -galactosidase gene in the E1 region and also contains the PD1 mutation in the penton gene. The PD1 mutation results in a substitution of

15 amino acids 337 through 344 of the penton protein, HAIRGDTF (SEQ ID No. 9), with amino acids SRGYPYDVPDYAGTS (SEQ ID No. 10), thus replacing the RGD tripeptide (see, Einfeld *et al.* (2001) *J. Virology* 75:11284-11291). The mutation in the penton gene was generated in the plasmid pGEMpen5, which contains the Adenovirus serotype 5 penton gene. To generate the mutation,

20 four oligonucleotides were synthesized. The sequences of the oligonucleotides were as follows: penton 1: 5' CGC GGA AGA GAA CTC CAA CGC GGC AGC CGC GGC AAT GCA GCC GGT GGA GGA CAT GAA 3' (SEQ ID No. 11); penton 2: 5' TAT CGT TCA TGT CCT CCA CCG GCT GCA TTG CCG CGG CTG CCG CGT TGG AGT TCT CTT CC 3' (SEQ ID No. 12); penton 3: 5' CGA TAG CCG

25 CGG CTA CCC CTA CGA CGT GCC CGA CTA CGC GGG CAC CAG CGC CAC ACG GGC TGA GGA GAA GCG CGC 3' (SEQ ID No. 13); penton 4: 5' TCA GCG CGC TTC TCC TCA GCC CGT GTG GCG CTG GTG CCC GCG TAG TCG GGC ACG TCG TAG GGG TAG CCG CGG C 3' (SEQ ID No. 14). The complementary oligonucleotides penton 1 and penton 2 were annealed to each other as were

30 penton 3 and penton 4. The duplex generated by annealing penton 3 and penton 4 encoded the substitution of amino acids 337 through 344 described above. The duplex generated by annealing penton 1 and penton 2 possessed a 5 base

-58-

- 5' overhang which was compatible to a 5 base 5' overhang on the duplex generated by annealing penton 3 and penton 4. The opposite end of the duplex generated by annealing penton 1 and penton 2 contained an Earl compatible overhang. The opposite end of the duplex generated by annealing penton 3 and penton 4 contained a BbvCI compatible overhang. The two duplexes were ligated to each other and ligated back into the pGEMpen5 backbone as follows. First, pGEMpen5 was digested with BbvCI and PstI and the resulting DNA fragments were separated by electrophoresis on an agarose gel. The 3360 bp fragment was excised from the gel and purified. The plasmid pGEMpen5 was also digested with PstI and Earl and the resulting fragments were separated by electrophoresis on an agarose gel. The 955 bp fragment was excised from the gel and purified. These two fragments from the pGEMpen5 plasmid were ligated with the two pairs of annealed oligonucleotides to generate the plasmid pGEMpen5PD1.
- 15 The mutated penton gene was transferred from pGEMpen5PD1 to pSQ1 using a 5-way ligation as follows. First, the region of the penton gene containing the PD1 mutation was excised from pGEMpen5PD1 by digestion with PvuI and AscI. The 974 bp fragment containing the PD1 mutation was purified. Four DNA fragments were prepared from the pSQ1 plasmid (Figure 3B) as follows. The plasmid was digested with Csp45I and FseI and the 9465 bp fragment was purified. In addition pSQ1 was digested with FseI and PvuI and the 2126 bp fragment was purified. The plasmid pSQ1 was digested with AscI and BamHI and the 5891 bp fragment was purified. Finally, pSQ1 was digested with BamHI and Csp45I and the 14610 bp fragment was purified. The 5 purified DNA fragments were ligated to each other to form the plasmid pSQ1PD1 (Figure 4).
- 20 To generate adenoviral vector, pSQ1PD1 was linearized by digestion with Clal and co-transfected into PerC6 cells with pAdmireRSVnBg (Figure 3A) which had been digested with SalI and PacI. hexadimethrine bromide was maintained in the medium at 4 μ g/ml. When a cytopathic effect was observed, a crude viral lysate was further expanded on PerC6 cells. The virus was purified by standard CsCl centrifugation procedures.
- 30

-59-

Av1nBgFKO1PD1

Genetic incorporation of the fiber KO1 or KO12 mutation in combination with the penton PD1 mutation to generate Av1nBgFKO1PD1

The adenoviral vectors Av1nBgFKO1PD1 and Av1nBgKO12PD1 were generated in an E1-, E3-deleted adenovirus serotype 5 genome. Both vectors contains a RSV promoted nuclear-localizing β -galactosidase gene in the E1 region and also contains either the KO1 or KO12 mutation in the fiber gene as well as the PD1 mutation in the penton gene. The vectors were constructed as follows. First, the plasmid pSQ1PD1 was digested with Csp45I and SpeI and the 23976 bp fragment containing the PD1 mutated penton gene was purified. In addition, the plasmids pSQ1KO1 or pSQ1KO12 (Figure 3B) were digested with Csp45I and SpeI and the 9090 bp fragment containing the KO1 or KO12 mutated fiber gene were purified. The appropriate purified fragments were ligated to each other to form the plasmid pSQ1FKO1PD1 (Figure 5A) or pSQ1KO12PD1 (Figure 5B) that contains the KO1 (or KO12) mutated fiber gene and the PD1 mutated penton gene. To generate virus, pSQ1FKO1PD1 or pSQ1KO12PD1 was linearized with ClaI and co-transfected into 633 cells with pAdmireRSVnBg (Figure 3A) which had been digested with Sall and PacI. After three rounds of amplification in 633 cells a cytopathic effect was observed and the crude viral lysate was then amplified on PerC6 cells. Hexadimethrine bromide was maintained in the medium at 4 μ g/ml. Each virus was purified by standard CsCl centrifugation procedures.

EXAMPLE 2***In Vitro* Evaluation of Adenoviral Vectors Containing the KO1 and PD1 Mutations**

Several recombinant adenoviral vectors were used in these studies to demonstrate the function of the KO1 fiber mutation and included Av1nBg, Av1nBgFKO1, Av1nBgPD1, and Av1nBgFKO1PD1, described above. The transduction efficiencies of adenoviral vectors containing the KO1 and/or PD1 mutations were evaluated on cells of the alveolar epithelial cell line A549. The transduction efficiencies were compared to that of Av1nBg, an adenoviral vector containing wild type fiber and penton.

-60-

The day prior to infection, cells were seeded into 24-well plates at a density of approximately 1×10^5 cells per well. Immediately prior to infection, the exact number of cells per well was determined by counting a representative well of cells. Each of the vectors, Av1nBg, Av1nBgFKO1, and Av1nBgFKO1PD1 were used to transduce A549 cells at each of the following particle per cell (PPC) ratios: 100, 500, 1000, 2500, 5000, 10,000. The cell monolayers were stained with X-gal 24 hours after infection and the percentage of cells expressing β -galactosidase was determined by microscopic observation and counting of cells. Transductions were done in triplicate and three random fields in each well were counted, for a total of nine fields per vector.

The results at the 500 PPC ratio are shown in Figure 6 and show a significantly reduced transduction efficiency on A549 cells using vectors containing the KO1 mutation alone or when combined with PD1 compared to Av1nBg. The vectors containing the PD1 mutation alone had no effect on adenoviral transduction of A549 cells *in vitro*.

EXAMPLE 3

In Vivo Analysis of Adenoviral Vectors Containing the FKO1 and PD1 Mutations

This Example provides experiments that evaluate the *in vivo* biodistribution of adenoviral vectors containing the KO1 and PD1 mutations and their influence on adenoviral-mediated liver transduction. The results show that ablating the viral interaction with CAR and/or integrins is not sufficient to fully detarget adenoviral vectors from the liver *in vivo*.

A positive control cohort received Av1nBg and a negative control group received HBSS. Additionally, the Av1nBgFKO12 and Av1nBgFKO12PD1 vectors were analyzed *in vivo*. These vectors each contain a fiber protein with the four amino acid substitution in the AB loop. Additionally, Av1nBgFKO12PD1 contains a mutation in the penton base. Both of these mutations were known (see, Einfeld *et al.* (2001) *J. Virology* 75:11284-11291), and were alleged to decrease liver transduction 10 to 700 fold, respectively. Cohorts of five C57BL/6 mice received each vector via tail vein injection at a dose of 1×10^{13} particles per kg. The animals were sacrificed approximately 72 hours after vector administration by carbon dioxide asphyxiation. Liver, heart, lung, spleen, and

-61-

kidney were collected from each animal. The median lobe of the liver was placed in neutral buffered formalin to preserve the sample for β -galactosidase immunohistochemistry. In addition, tissue from each organ was frozen to preserve it for hexon PCR analysis to determine vector content. A separate
5 sample of liver from each mouse was frozen to preserve it for a chemiluminescent β -galactosidase activity assay.

For β -galactosidase immunohistochemistry slices of liver, approximately 2-3 mm thick, were placed in 10% neutral buffered formalin. After fixation, these samples were embedded in paraffin, sectioned, and analyzed by
10 immunohistochemistry for β -galactosidase expression. A 1:1200 dilution was used of a rabbit anti- β -galactosidase antibody (ICN Pharmaceuticals, Inc.; Costa Mesa, CA) in conjunction with a Vectastain ABC kit (Vector Laboratories, Inc., Burlingame, CA) to visualize positive cells.

The chemiluminescent β -galactosidase activity assay was performed
15 using the Galacto-Light Plus™ chemiluminescent assay (Tropix, Inc., Foster City, CA) system. Tissue samples were collected in lysis matrix tubes containing two ceramic spheres (Bio101, Carlsbad, CA) and frozen on dry ice. The tissues were thawed and 500 μ l of lysis buffer from the Galacto-Light Plus kit was added to each tube. The tissue was homogenized for 30 seconds using a
20 FastPrep System (Bio101, Carlsbad, CA). Liver samples were homogenized for an additional 30 seconds. β -galactosidase activity was determined in the liver homogenates according to the manufacture's protocol.

For hexon PCR analysis DNA from tissues was isolated using the Qiagen Blood and Cell Culture DNA Midi or Mini Kits (Qiagen Inc., Chatsworth, CA).
25 Frozen tissues were partially thawed and minced using sterile disposable scalpels. Tissues were then lysed by incubation overnight at 55° C in Qiagen buffer G2 containing 0.2 mg/ml RNaseA and 0.1 mg/ml protease. Lysates were vortexed briefly and then applied to Qiagen-tip 100 or Qiagen-tip 25 columns. Columns were washed and DNAs were eluted as described in the manufacturer's
30 instructions. After precipitation, DNAs were dissolved in water and the concentrations were spectrophotometrically determined (A260 and A280) on a

-62-

DU-600 (Beckman Coulter, Inc.; Fullerton, CA) or a SPECTRAMax PLUS (Molecular Devices, Inc.; Sunnyvale, CA) spectrophotometer. 2.3.2.

PCR primers and a Taqman probe specific to adenovirus hexon sequences were designed using Primer Express software v. 1.0 (Applied Biosystems, Foster

5 City, CA). Primer and probe sequences were:

Hexon Forward primer: 5'-CTTCGATGATGCCGAGTG-3' (SEQ ID No. 38);

Hexon Reverse primer: 5'-GGGCTCAGGTACTCCGAGG-3' (SEQ ID No. 39); and

Hexon Probe: 5'-FAM-TTACATGCACATCTCGGGCCAGGAC-TAMRA-3' (SEQ ID No. 40).

10 Amplification was performed in a reaction volume of 50 μ l under the following conditions: 10 ng (tumor) or 1 μ g (liver and lung) of sample DNA, 1X Taqman Universal PCR Master Mix (Applied Biosystems), 600 nM forward primer, 900 nM reverse primer and 100 nM hexon probe. Thermal cycling conditions were: 2 minute incubation at 50° C, 10 minutes at 95° C, followed by

15 35 cycles of successive incubation at 95° C for 15 seconds and 60° C for 1 minute. Data was collected and analyzed using the 7700 Sequence Detection System software v. 1.6.3 (Applied Biosystems). Quantification of adenovirus copy number was performed using a standard curve that includes dilutions of adenovirus DNA from 1,500,000 copies to 15 copies in the appropriate

20 background of cellular genomic DNA. For analysis of tumor tissues, a standard curve in a background of 10 ng human DNA was generated. For analysis of mouse liver and lung tissues, a standard curve using the same adenovirus DNA dilutions in a background of 1 μ g CD-1 mouse genomic DNA was generated. Samples were amplified in triplicate, and the average number of total copies was

25 normalized to copies per cell based on the input DNA weight amount and a genome size of 6×10^9 bp.

The results of the β -galactosidase activity assay and adenoviral hexon DNA content for liver transduction by these vectors are shown in Figure 7A and 7B. The vector containing the KO1 or KO12 mutations alone showed, on

30 average, a slight increase in liver transduction compared to Av1nBg, which is consistent with several previous experiments. The vectors containing the PD1 mutation alone or combined with KO1 or KO12 showed a slight decrease in liver

-63-

transduction compared to Av1nBg, suggesting that integrins are involved to some extent in hepatic uptake of the adenoviral vectors.

The results of the immunohistochemical staining of liver sections for β -galactosidase were consistent with the activity assays (data not shown) and demonstrate that gene expression was localized specifically to hepatocytes. The vectors containing the KO1 or KO12 mutation alone showed a slight increase in liver transduction as revealed by a more intense and frequent immunohistochemical-staining pattern. The vectors containing the PD1 mutation, either alone or combined with KO1 or KO12, showed little difference in transduction compared to Av1nBg. These results demonstrate that ablating the viral interaction with CAR and/or integrins is not sufficient to fully detarget adenoviral vectors from the liver *in vivo*.

In summary, the fiber AB loop mutation contained in Av1nBgFKO1 or Av1nBgKO12 ablates interaction with human and mouse CAR *in vitro* and diminished transduction *in vitro*. *In vivo*, however, fiber AB loop mutations behaved unexpectedly, because such mutations were found to enhance adenoviral-mediated gene transfer to liver and results in increasing vector potency. The penton base, PD1 mutation that ablates interaction with the second receptor involved in adenoviral internalization had no effect *in vitro* and little to no effect *in vivo*. These studies indicated that other receptors are responsible for adenoviral gene transfer to the liver *in vivo*.

EXAMPLE 4

Description Of Adenoviral Vectors Containing A Fiber With Amino Acid Substitutions At The Heparin Sulfate Binding Domain In The Fiber Shaft

Vectors containing substitutions at all four of the amino acids in the four amino acid motif in the Ad5 fiber shaft (residues 91 to 94, KKTK; SEQ ID No. 1) were generated in order to ablate the potential interaction with HSP. The mutation is termed HSP because it potentially eliminates binding to heparan sulfate proteoglycans. Vectors containing the HSP mutation alone and combined with the KO1 mutation (fiber knob AB loop mutation that ablates CAR binding), the PD1 mutation (penton mutation that eliminates RGD/integrin interaction), and a triple knockout vector (HSP, KO1, PD1) were generated.

-64-

Generation of the HSP fiber mutation: The HSP mutation was incorporated into the fiber gene by using a PCR-based strategy of gene splicing by overlap extension (PCR SOEing). First, a segment of the Ad5 genome extending from within the E3 region into the 5' end of the fiber gene was amplified by PCR using the plasmid pSQ1 (Figure 3B) as a template and two primers termed 5FF and 5HSPR. The DNA sequence of 5FF is as follows: 5' GAA CAG GAG GTG AGC TTA GA 3' (SEQ ID No. 5). This sequence corresponds to base pairs 25,199 - 25,218 of pSQ1. The DNA sequence of 5HSPR is as follows: 5' GGC TCC GGC TCC GAG AGG TGG GCT CAC AGT GGT TAC ATT T 3' (SEQ ID No. 15). 5HSPR is a reverse primer for 5FF and corresponds to a region in the fiber shaft adjacent to the KKTK (SEQ ID No. 1) region. The primer contains a 5' extension that encodes a GAGA substitution for the native KKTK (encoded by SEQ ID No. 1) amino acid sequence. A second PCR using pSQ1 as a template amplified the region immediately 3' of the KKTK (SEQ ID No. 1) site and extending past the MunI site located 40 base pairs 3' of the stop codon for the fiber gene. The two primers used for this reaction were 3HSPF and 3FR. The DNA sequence of 3HSPF is as follows: 5' GGA GCC GGA GCC TCA AAC ATA AAC CTG GAA AT 3' (SEQ ID No. 16). It contains a 5' extension that is complementary to the 5' extension of 5HSPR. The DNA sequence of 3FR is as follows: 5' GTG GCA GGT TGA ATA CTA GG 3' (SEQ ID No. 8).

The two PCR products were joined by PCR SOEing using primers 5FF and 3FR. The resulting PCR product was digested with the restriction enzymes XbaI and MunI. The 2355 bp fragment was gel purified and ligated with the 6477 bp XbaI to MunI fragment of the plasmid pFBshuttle(EcoRI) (Figure 8) to generate the plasmid pFBSEHSP. The plasmid pFBshuttle(EcoRI) was generated by digesting the plasmid pSQ1 with EcoRI, then gel purifying and self-ligating the 8.8 kb fragment containing the fiber gene. Next, the fiber gene containing the HSP mutation was transferred from pFBSEHSP into pSQ1 using a three-way ligation. The 16,431 bp EcoRI to NdeI fragment of pSQ1, the 9043 bp NdeI to XbaI fragment of pSQ1, and the 7571 bp XbaI to EcoRI fragment of pFBSEHSP were isolated and ligated to generate pSQ1HSP (Figure 9).

-65-

To generate a recombinant adenoviral vector containing the HSP mutation in the fiber gene, pSQ1HSP was digested with ClaI and pAdmireRSVnBg (Figure 3A) was digested with Sall and PacI, then the two digested plasmids were co-transfected into 633 cells (von Seggern *et al.* (2000) *J Virology* 74:354-362).

- 5 Homologous recombination between the two plasmids generated a full-length adenoviral genome capable of replication in 633 cells, which inducibly express Ad5 E1A and constitutively express wild-type fiber protein. After propagation on 633 cells, the virus capsid contained wildtype and mutant fiber proteins. To obtain viral particles containing only the modified fiber with the HSP mutation, the viral preparation was used to infect PerC6 cells, which do not express fiber. The resulting virus, termed Av1nBgFS*, was purified by standard CsCl centrifugation procedures.

Generation of vector containing the HSP and KO1 mutations

- To generate an adenoviral vector containing the HSP and KO1 mutations in fiber, a PCR SOEing strategy identical to the one described above was used except that the plasmid pSQ1FKO1 was used as the template. The PCR SOEing product was digested with XbaI and MunI and ligated with the 6477 bp XbaI to MunI fragment of pFBshuttle(EcoRI) to generate pFBSEHSPKO1. The fiber gene containing the HSP and KO1 mutations was transferred from pFBSEHSPKO1 into the pSQ1 backbone using a three-way ligation strategy identical to the one described above for the HSP mutation alone, to generate the plasmid pSQ1HSPKO1 (Figure 10). Recombinant adenoviral vector containing the HSP and KO1 mutations in the fiber gene was generated by co-transfecting pSQ1HSPKO1 digested with ClaI and pAdmireRSVnBg digested with Sall and PacI into 633 cells. Adenovirus was propagated and purified as described above for the vector containing the HSP mutation alone. The resulting virus was termed Av1nBgFKO1S*.

-66-

Generation of vector containing the HSP and PD1 mutations

The following strategy was used to generate a recombinant adenoviral vector containing the fiber HSP mutation and the penton PD1 mutation. The plasmid pSQ1PD1 (Figure 4) was digested with the restriction enzymes Csp45I and SpeI and the 23,976 bp fragment was isolated and purified. In addition, the plasmid pSQ1HSP was also digested with Csp45I and SpeI and the 9090 bp fragment was isolated and purified and ligated to the 23,976 bp fragment to generate the plasmid pSQ1HSPPD1 (Figure 11), which contains the fiber HSP and penton PD1 mutations. An adenoviral vector was generated, propagated, and purified as described above. The resulting virus was termed Av1nBgS*PD1.

Generation of vector containing the HSP, KO1, and PD1 mutations

To generate an adenoviral vector containing the HSP, KO1, and PD1 mutations the following strategy was used. First, the plasmid pSQ1PD1 was digested with Csp45I and SpeI and the 23,976 bp fragment was isolated and purified. In addition, the plasmid pSQ1HSPKO1 was digested with Csp45I and SpeI and the 9090 bp fragment was isolated and purified. The two DNA fragments were ligated to form the plasmid pSQ1HSPKO1PD1 (Figure 12). Recombinant adenoviral vector was generated, propagated, and purified as described above. The resulting virus was termed Av1nBgFKO1S*PD1.

20

EXAMPLE 5

In Vitro Evaluation Of Adenoviral Vectors Containing The HSP Fiber Mutation

The transduction efficiencies of adenoviral vectors containing the HSP mutation in the fiber gene, either alone or combined with the KO1 and/or PD1 mutations, were evaluated on A549 and HeLa cells. The transduction efficiencies were compared to that of Av1nBg, an adenoviral vector containing wild type fiber and penton. The day prior to infection, cells were seeded into 24-well plates at a density of approximately 1×10^5 cells per well. Immediately prior to infection, the exact number of cells per well was determined by counting a representative well of cells. Each of the vectors, Av1nBg (see, Stevenson *et al.* (1997) *J. Virol.* 71:4782-4790), Av1nBgS*, Av1nBgFKO1S*, Av1nBgS*PD1, and Av1nBgFKO1S*PD1, were used to transduce A549 cells at each of the following particle per cell (PPC) ratios: 100, 500, 1000, 2500,

-67-

5000, 10,000. HeLa cells were transduced with each of the above vectors, as well as a vector containing the KO1 mutation alone (Av1nBgFKO1) and a vector containing the PD1 mutation alone (Av1nBgPD1) at 2000 PPC. The cell monolayers were stained with X-gal 24 hours after infection and the percentage of cells expressing β -galactosidase was determined by microscopic observation and counting of cells. Transductions were done in triplicate and three random fields in each well were counted, for a total of nine fields per vector.

The results (depicted in Figures 13A-13B) showed significantly reduced transduction efficiencies on A549 and HeLa cells using vectors containing the HSP mutation compared to Av1nBg. The vectors containing the HSP mutations, however, demonstrated a dose response on A549 cells, in that increasing PPC ratios yielded increasing transduction.

Competition experiments were done to determine which receptor molecular interactions are involved in transduction of A549 cells by the various vectors.

Transductions were performed in the presence or absence of various competitors including Ad5 fiber knob, a 50 amino acid oligopeptide derived from Adenovirus serotype 2 penton base which spans the RGD tripeptide region, or heparin (Invitrogen Life Technologies, Gaithersburg, MD). Monolayers of A549 cells were cultured in Richters medium supplemented with 10% FBS and were transduced with Av1nBg, Av1nBgS*, Av1nBgFKO1S*, Av1nBgS*PD1, or Av1nBgFKO1S*PD1 in infection medium (IM, Richters medium plus 2% FBS). Different PPC ratios were used for the different vectors to achieve measurable transduction levels. The PPC ratios were as follows: Av1nBg: 500 PPC, Av1nBgS*: 10,000 PPC, Av1nBgFKO1S*: 20,000 PPC, Av1nBgS*PD1: 10,000 PPC, and Av1nBgFKO1S*PD1: 20,000 PPC. Fiber knob competition was performed by pre-incubating cells in IM containing 16 μ g/ml of fiber knob for 10 minutes at room temperature prior to infection with virus. Penton base peptide competition was performed by pre-incubating cells in IM containing 500nM peptide for 10 minutes at room temperature prior to infection with virus.

Heparin competition was performed by pre-incubating each adenoviral vector in IM containing 3 mg/ml of heparin for 20 minutes at room temperature. In all cases, the competitor remained in the IM during the 1 hour infection when virus

-68-

was rocked on the cell monolayers at 37° C in 5% CO₂. After infection, the monolayers were washed with PBS, 1 ml of complete medium was added per well and the cells were incubated for an additional 24 hours to allow for β -galactosidase expression. The cell monolayers were then fixed and stained with X-Gal. The percentage of cells transduced was determined by light microscopy as described above. Each condition was carried out in triplicate and three random fields per well were counted, for a total of nine fields per condition. The average percentage of transduction per high-power field was determined.

The results of the competition experiment (Figure 13C) showed that fiber knob inhibited transduction of cells by all vectors except for those that contained the KO1 mutation. The penton base peptide only inhibited transduction by Av1nBgFKO1S*. Heparin inhibited transduction by Av1nBgFKO1S* and Av1nBgFKO1S*PD1, but did not affect transduction by any of the other viruses suggesting the presence of additional heparin binding sites on the adenoviral capsid but that the shaft contains the predominant site.

EXAMPLE 6

In Vivo Analysis Of Adenoviral Vectors Containing The HSP Mutation In Fiber

The objective of this study was to evaluate the *in vivo* biodistribution of adenoviral vectors containing the HSP mutation and to determine whether this shaft modification influences adenoviral-mediated liver transduction. In addition, vectors containing the HSP mutation combined with KO1, or PD1, or a combination of all three mutations were evaluated as well as vectors containing the KO1 mutation alone and the PD1 mutation alone. A positive control cohort received Av1nBg and a negative control group received HBSS. Cohorts of five C57BL/6 mice received each vector via tail vein injection at a dose of 1×10^{13} particles per kg. The animals were sacrificed approximately 72 hours after vector administration by carbon dioxide asphyxiation. Liver, heart, lung, spleen, and kidney were collected from each animal. The median lobe of the liver was placed in neutral buffered formalin to preserve the sample for β -galactosidase immunohistochemistry. In addition, tissue from each organ was frozen to preserve it for hexon real time PCR analysis to determine vector content. A

-69-

separate sample of liver from each mouse was frozen to preserve it for a chemiluminescent β -galactosidase activity assay. β -galactosidase immunohistochemistry, hexon real-time PCR and the chemiluminescent β -galactosidase activity assay were carried out as described in Example 3.

5 The results of the β -galactosidase activity assay (Figure 14A) and adenoviral hexon DNA content (Figure 14B) showed a dramatic reduction in liver transduction by vectors containing the HSP mutation. The vectors containing the HSP mutation alone resulted in reducing adenoviral-mediated liver gene expression by approximately 20-fold. When combined with the KO1 mutation
10 (HSP, KO1, PD1), yielded approximately a 1000-fold reduction in β -galactosidase activity in the liver compared to the control vector Av1nBg. The vector containing the KO1 mutation alone showed a slight increase, on average, in liver transduction compared to Av1nBg, which is consistent with several previous experiments. The vectors containing the PD1 mutation alone or combined with
15 KO1 showed a slight decrease in liver transduction compared to Av1nBg, although the decrease was not statistically significant. Analysis of hepatic adenoviral hexon DNA content (Figure 14B) confirmed these results.

The results of the immunohistochemical staining of liver sections for β -galactosidase were consistent with the activity assays (data not shown) and
20 demonstrated that gene expression was localized specifically to hepatocytes. Vectors containing the HSP mutation, either alone or in combination with KO1 and/or PD1, showed a dramatic reduction in hepatocyte transduction. The vector containing the KO1 mutation alone showed a slight increase in liver transduction as revealed by a more intense and frequent immunohistochemical
25 staining pattern. The vectors containing the PD1 mutation, either alone or combined with KO1, showed little difference in transduction compared to Av1nBg.

EXAMPLE 7

30 Description of Adenoviral Vectors Containing the HSP Fiber Shaft Mutation with and without the KO1 Fiber Mutation and with and without a cRGD Targeting Ligand in the Fiber Knob HI Loop

-70-

Generation of vector containing the HSP fiber shaft mutation and a cRGD ligand in the HI loop: The following strategy was used to generate an adenoviral vector containing a fiber with the HSP shaft mutation and a cRGD ligand in the HI loop. The plasmid p5FloxHRFRGD was digested with the restriction enzymes BstXI and KpnI and the 1157 bp fragment was isolated and purified. In addition, the fiber shuttle plasmid pFBSEHSP, described in Example 1 above, was digested with BstXI and KpnI and the 4549 bp and 3156 bp fragments were isolated and purified. The three fragments were ligated to generate the plasmid pFBSEHSPRGD, which encodes a fiber containing the HSP mutation and cRGD in the HI loop. The fiber gene from this plasmid was transferred into the pSQ1 backbone as follows. The plasmid pFBSEHSPRGD was digested with EcoRI and XbaI and the 7601 bp fragment was isolated and purified. The plasmid pSQ1 (Figure 3B) was digested with the restriction enzymes EcoRI, NdeI, and XbaI and the 16,431 bp EcoRI to NdeI fragment and the 9043 bp NdeI to XbaI fragment were isolated and purified. The three DNA fragments were ligated to generate the plasmid pSQ1HSPRGD (Figure 15A).

To generate a recombinant adenoviral vector containing the HSP mutation in the fiber gene along with a cRGD ligand in the HI loop, the plasmid pSQ1HSPRGD was digested with ClaI and co-transfected into 633 cells with pAdmireRSVnBg which had been digested with SalI and PacI. After propagation on 633 cells, the virus capsid contained wildtype and mutant fiber proteins. To obtain viral particles containing only the modified fiber with the HSP mutation and a cRGD ligand, the viral preparation was used to infect PerC6 cells, which do not express fiber. The resulting virus, termed Av1nBgS*RGD, was purified by standard CsCl centrifugation procedures.

Generation of vector containing the HSP fiber shaft mutation, the KO1 fiber knob mutation, and a cRGD ligand in the HI loop

The following strategy was used to generate an adenoviral vector containing a fiber with the HSP shaft mutation, the KO1 fiber knob mutation, and a cRGD ligand in the HI loop. The plasmid p5FloxHRFRGD was digested with the restriction enzymes BstXI and KpnI and the 1157 bp fragment was isolated and purified. In addition, the fiber shuttle plasmid pFBSEHSPKO1,

-71-

described in Example 1 above, was digested with BstXI and KpnI and the 4549 bp and 3156 bp fragments were isolated and purified. The three fragments were ligated to generate the plasmid pFBSEHSPKO1RGD, which encodes a fiber containing the HSP mutation, the KO1 mutation, and cRGD in the HI loop. The fiber gene from this plasmid was transferred into the pSQ1 backbone as follows. The plasmid pFBSEHSPKO1RGD was digested with EcoRI and XbaI and the 7601 bp fragment was isolated and purified. The plasmid pSQ1 (Figure 3B) was digested with the restriction enzymes EcoRI, NdeI, and XbaI and the 16,431 bp EcoRI to NdeI fragment and the 9043 bp NdeI to XbaI fragment were isolated and purified. The three DNA fragments were ligated to generate the plasmid pSQ1HSPKO1RGD (Figure 15B).

To generate a recombinant adenoviral vector containing the HSP and KO1 mutations in the fiber gene along with a cRGD ligand in the HI loop, the plasmid pSQ1HSPKO1RGD was digested with ClaI and co-transfected into 633 cells with pAdmireRSVnBg which had been digested with Sall and PacI. After propagation on 633 cells, the virus capsid contained wildtype and mutant fiber proteins. To obtain viral particles containing only the modified fiber with the HSP and KO1 mutations and a cRGD ligand, the viral preparation was used to infect PerC6 cells, which do not express fiber. The resulting virus, termed Av1nBgFKO1S*RGD, was purified by standard CsCl centrifugation procedures.

EXAMPLE 8

***In Vitro* Evaluation of Adenoviral Vectors Containing the HSP Fiber Shaft Mutation with or without the Fiber Knob KO1 Mutation and with or without a cRGD Ligand in the HI Loop**

The transduction efficiencies of adenoviral vectors containing the HSP fiber shaft mutation with or without the fiber KO1 mutation and with or without the cRGD ligand in the HI loop were evaluated on A549 cells. The transduction efficiencies were compared to that of Av1nBg, an adenoviral vector containing wild type fiber. The day prior to infection, cells were seeded into 24-well plates at a density of approximately 1×10^5 cells per well. Immediately prior to infection, the exact number of cells per well was determined by counting a representative well of cells. Each of the vectors, Av1nBg, Av1nBgS*,

-72-

Av1nBgFKO1S*, Av1nBgS*RGD, and Av1nBgFKO1S*RGD, were used to transduce A549 cells at a particle to cell ratio of 6250. The cell monolayers were stained with X-gal 24 hours after infection and the percentage of cells expressing β -galactosidase was determined by microscopic observation and counting of cells. Transductions were done in triplicate and three random fields in each well were counted, for a total of nine fields per vector. The results (Figure 16) showed that the cRGD ligand dramatically increased the transduction efficiencies of vectors containing the HSP mutation alone or combined with the KO1 mutation. Av1nBgS* yielded approximately 22% positive cells, while Av1nBgS*RGD yielded approximately 95% positive cells. Similarly, Av1nBgFKO1S* yielded only 4% positive cells, while Av1nBgFKO1S*RGD yielded 85% positive cells. Therefore, the vector containing the shaft mutation is viable and can be retargeted with the addition of a ligand.

EXAMPLE 9

15 Construction Of Ad5 Vectors Containing The Ad35 Fiber And Derivatives Thereof

The KO1 and HSP mutations in the Ad5 fiber protein (5F), described above, were designed to ablate interactions that are responsible for the normal tropism of the Ad5 virus. An alternative strategy to detarget the virus is to replace the Ad5 fiber with a fiber from another serotype which does not bind CAR and which does not possess the heparin sulfate proteoglycan (HSP) binding domain (KKTK; SEQ ID No. 1) within the shaft. The fiber of adenovirus serotype 35 (35F) does not bind CAR and does not possess the HSP binding domain in its shaft. Replacement of the 5F with the 35F can detarget the liver and provide a suitable platform for retargeting the vector to the desired tissue.

Generation of an Ad5 based vector containing the Ad35 fiber: A PCR SOEing strategy was used to generate a vector based on the Ad5 serotype but containing the Ad35 fiber in place of the Ad5 fiber. First, PCR was used to amplify a region in the plasmid pSQ1 between the XbaI site at bp 25,309 and the start of the fiber gene. The primers used for this reaction were P-0005/U and P-0006/L. The DNA sequence of P-0005/U was as follows: 5' C TCT AGA AAT GGA CGG AAT TAT TAC AG 3' (SEQ ID No. 17). This sequence

-73-

corresponds to bp 25,308 through 25,334 of pSQ1. The DNA sequence of P-0006/L was as follows: 5' TCT TGG TCA TCT GCA ACA ACA TGA AGA TAG TG 3' (SEQ ID No. 18). It contains a 10 base pair 5' extension that is complementary to the start of the Ad35 fiber gene, while the remainder of the primer anneals to the sequence immediately 5' of the ATG start codon of the fiber gene in pSQ1. A PCR product of the expected size, 583 bp, was obtained and the DNA was gel purified. A second PCR amplified the Ad35 fiber gene using DNA extracted from wildtype Ad35 virus as a template. The primers used for this reaction were P-0007/U and 35FMun. The DNA sequence of P-0007/U was as follows: 5' GT TGT TGC AG ATG ACC AAG AGA GTC CGG CTC A 3' (SEQ ID No. 19). It contains a 10 base pair 5' extension that is homologous to the 10 bp immediately prior to the ATG start codon of the fiber gene in Ad5. The remainder of the primer anneals to the start of the Ad35 fiber gene. The DNA sequence of 35FMun was as follows: 5' AG CAA TTG AAA AAT AAA CAC GTT GAA ACA TAA CAC AAA CGA TTC TTT A GTT GTC GTC TTC TGT AAT GTA AGA A 3' (SEQ ID No. 20). It contains a 46 base pair 5' extension that is complementary to the region of the Ad5 genome between the end of fiber and the MunI site 40 bp downstream of the fiber gene. In addition, the 5' extension encodes the last amino acid and stop codon of the Ad5 fiber gene. This region was retained in the vector because it contains the polyadenylation site for the fiber gene. The remainder of the primer anneals to the 3' end of the Ad35 fiber gene, up to the next to last amino acid codon. A PCR product of the expected size, 1027 bp, was obtained and the DNA was gel purified. The two PCR products were mixed and joined together by PCR SOEing using primers P-0005/U and P-0009. The DNA sequence of P-0009 was as follows: 5' AG CAA TTG AAA AAT AAA CAC GTT G 3' (SEQ ID No. 21). It corresponds to bp 27,648 through 27,669 of pSQ1 and overlaps the MunI site in that region. A PCR product of the expected size, 1590 bp, was obtained and gel purified. It was cloned into the plasmid pCR4blunt-TOPO (Invitrogen Corporation, Carlsbad CA) using the Zero Blunt TOPO PCR Cloning Kit from Invitrogen. This intermediate cloning step simplified DNA sequencing of the PCR SOEing product. The resulting plasmid, termed pTOPOAd35F, was digested with XbaI and MunI

-74-

and the 1585 bp digestion product was gel purified and ligated with the 6477 bp fragment of pFBshuttle(EcoRI) digested with XbaI and MunI to generate the plasmid pFBshuttleAd35F. The Ad35 fiber gene was transferred from pFBshuttleAd35F into pSQ1 as follows. The plasmid pSQ1 was digested with
5 EcoRI and the 24,213 bp fragment was gel purified. The plasmid pFBshuttleAd35F was linearized with EcoRI and ligated with the 24,213 bp fragment from pSQ1. Restriction diagnostics were performed to screen for constructs containing the Ad35 fiber gene inserted into the pSQ1 backbone in the correct orientation. The pSQ1 plasmid containing the Ad35 fiber gene in the
10 proper orientation was termed pSQ1Ad35Fiber (Figure 17A). To generate adenoviral vector containing the Ad35 fiber, pSQ1Ad35Fiber was digested with ClaI and co-transfected into 633 cells with pAdmireRSVnBg which had been digested with SalI and PacI. After propagation on 633 cells, the resulting virus contained Ad5 fiber and Ad35 fibers on its capsid. The virus was amplified on
15 PerC6 cells to generate virus containing only the Ad35 fiber on its capsid. The resulting virus preparation was termed Av1nBg35F.

Construction of adenoviral vectors containing chimeric fibers derived from Ad5 and Ad35: Two chimeric fiber constructs were prepared by PCR gene overlap extension using plasmids containing the full length Ad5 or Ad35 fiber
20 cDNAs as templates. The Ad5 fiber tail and shaft regions (5TS; amino acids 1 to 403) were connected with the Ad35 fiber head region (35H; amino acids 137 to 323) to form the 5TS35H chimera, and the Ad35 fiber tail and shaft regions (35TS; amino acids 1 to 136) were connected with the Ad5 fiber head region (5H; amino acids 404 to 581) to form the 35TS5H chimera. The fusions were
25 made at the conserved TLWT sequence at the fiber shaft-head junction.

For the construction of the 5TS35H chimera, the pFBshuttle(EcoRI) plasmid was used as the template with primers P1 and P2 to generate the 5' fragment. The 3' fragment was generated using the pFBshuttleAd35 plasmid as the template with the P3 and P4 primers. The sequence of each primer used in
30 the construction of these chimeric fibers is listed in Table 2. Amplified PCR products of the expected size were obtained and were gel purified. A second PCR was carried out with the end primers P1 and P4 to join the two fragments

-75-

together. The DNA fragment generated in the second PCR was digested with Xba1 and Mun1 and was cloned directly into pFBshuttle(EcoR1) to create the fiber shuttle plasmid pFBshuttle5TS35H.

TABLE 2

5 Primers Used For The Exchange Of Fiber Shaft Regions Between Ad5 And Ad35 Fibers

Primer designation	Sequence	SEQ ID
P1	5'-GAACAGGAGGTGAGCTTAGA-3'	22
P2	5'-GTTAGGTGGAGGGTTTATTCCGGTCCAC AAAGTTAGCTTATC-3'	23
P3	5'-GATAAGCTAACTTTGTGGACCGGAATAAA CCCTCCACCTAAC-3'	24
P4	5'-GTGGCAGGTTGAATACTAGG-3	25
P5	5'-GTTAGGAGATGGAGCTGGTGTAGTCCATA AGGTGTTAATAC-3'	26
P6	5'-GTATTAACACCTTATGGACTACACCAGCT CCATCTCCTAAC-3'	27
P7	5'-TGCGCAAAAACAATCACCACGACAATCACAAT GTACATTGGAAGAAATCATACG-3'	28
P8	5'-ACATTGTGATTGTCGTGGTGATT GTTTTTGCGCATATGCCATACAATTTGAATG-3'	29

For the construction of the 35TS5H chimera, the pFBshuttleAd35 plasmid was used as the template with the P1 and P5 primers to generate the 5' fragment. The 3' fragment was generated using the pFBshuttle(EcoR1) plasmid as the template with the P6 and P4 primers. Following the same procedure described above, the fiber shuttle plasmid pFBshuttle35TS5H was generated.

For the 35TS5H and 5TS35H chimeras, the fiber gene was transferred from the pFBshuttle(EcoRI) backbone into pSQ1 as described above for the vector containing the Ad35 fiber. The resulting plasmids were called pSQ135T5H (Figure 18A) and pSQ15T35H (Figure 18B). In addition, adenoviral vectors were generated using the co-transfection strategy described above.

Construction of Ad5 vectors containing the Ad35 fiber with a cRGD targeting peptide in the HI loop of the 35F fiber knob: To incorporate the cRGD

-76-

targeting peptide into the Ad35 fiber HI loop, the P7 and P8 oligonucleotide primers encoding the ten amino acid sequence HCDCRGDCFC (SEQ ID No. 30) were synthesized. The pFBshuttleAd35 plasmid containing the full length Ad35 fiber cDNA was used as the template in the PCR reaction with the P1 and P7
5 primer pair or with the P4 and P8 primer pair in order to generate the 5' and 3' PCR fragments. A second PCR was then carried out with the end primers P1 and P4 to join the two fragments together. The resulting PCR fragment was digested with Xba1 and Mun1 and was cloned into pFBshuttle(EcoR1) to create the fiber shuttle plasmid pFBshuttleAd35cRGD. The modified Ad35 fiber gene
10 was transferred into pSQ1 using the EcoRI cloning strategy described above to generate pSQ1Ad35FcRGD (Figure 17B). Adenoviral vector was generated using the co-transfection strategy described above.

EXAMPLE 10

In Vitro Evaluation Of Adenoviral Vectors Containing 35F And Derivatives 15 Thereof

The transduction efficiencies of adenoviral vectors containing the 35F or derivatives thereof were evaluated on A549 cells. The transduction efficiencies were compared to that of Av1nBg, an adenoviral vector containing the 5F fiber. The day prior to infection, cells were seeded into 24-well plates at a density of
20 approximately 1×10^5 cells per well. Immediately prior to infection, the exact number of cells per well was determined by counting a representative well of cells. Each of the vectors, Av1nBg, Av1nBg35F, Av1nBg5T35H and Av1nBg35T5H were used to transduce A549 cells from 0 up to 1,000 particle per cell (PPC) ratios. The cell monolayers were stained with X-gal 24 hours after
25 infection and the percentage of cells expressing β -galactosidase was determined by microscopic observation and counting of cells. Transductions were done in triplicate and three random fields in each well were counted, for a total of nine fields per vector. The results (Figure 19) showed similar transduction efficiencies on A549 cells using the Av1nBg35F and Av1nBg5T35H vectors
30 compared to Av1nBg. The Av1nBg35T5H showed much lower transduction efficiencies on A549 cells compared to Av1nBg as a result of the Ad35 shaft domain. The Ad35 shaft domain does not contain a HSP binding motif and the

-77-

Av1nBg35T5H vector behaves similarly to the Av1nBgS* vector *in vitro* and *in vivo*. These studies also demonstrate that vectors containing fiber proteins without an HSP binding site are fully viable.

EXAMPLE 11

5 *in Vivo* Evaluation Of Adenoviral Vectors Containing 35F And Derivatives Thereof

The objective of this study was to evaluate the *in vivo* biodistribution of adenoviral vectors containing 35F fibers and derivatives thereof to determine whether vectors containing these fibers ablate liver transduction due to their shaft regions. A positive control cohort received Av1nBg and a negative control group received HBSS. Cohorts of five C57BL/6 mice received each vector via tail vein injection at a dose of 1×10^{13} particles per kg. The animals were sacrificed approximately 72 hours after vector administration by carbon dioxide asphyxiation. Liver, heart, lung, spleen, and kidney were collected from each animal. The median lobe of the liver was placed in neutral buffered formalin to preserve the sample for β -galactosidase immunohistochemistry. In addition, tissue from each organ was frozen to preserve it for hexon PCR analysis to determine vector content. A separate sample of liver from each mouse was frozen to preserve it for a chemiluminescent β -galactosidase activity assay. β -galactosidase immunohistochemistry, hexon real-time PCR and the chemiluminescent β -galactosidase activity assay were carried out as described in example 3.

The results of the β -galactosidase activity assay showed a dramatic reduction in liver transduction by vectors containing the Ad35 fiber or the 35T5H derivative (Figure 20) with an approximately 4- to 24-fold reduction in β -galactosidase activity in the liver compared to the control vector Av1nBg. These data demonstrate that shaft domains without HSP binding sites can effectively ablate hepatic *in vivo* gene transfer. In particular, HSP is the major entry mechanism for liver *in vivo*. CAR binding is a minor entry pathway.

-78-

EXAMPLE 12**Construction Of Ad5 Vectors Containing The Ad Serotype 41 Short Fiber And Derivatives Thereof**

The human adenovirus serotype 41 contains two different fibers on its capsid, encoded by two adjacent genes. One fiber has a molecular weight of 60kDa and is approximately 315A in length and is termed the long fiber. The other fiber has a molecular weight of 40kDa and is approximately 250+ in length and is termed the short fiber. The Ad41 short fiber does not bind CAR and does not possess the heparin binding domain (KKTK) in its shaft. Therefore, this fiber provides a useful platform for adenoviral vector targeting.

Construction of adenoviral vectors based on Ad5 but containing the Ad41 short fiber: A PCR SOEing strategy was used to generate a vector based on the Ad5 genome but containing the Ad41 short (Ad41s) fiber. First, PCR was used to amplify the region of pSQ1 between the XbaI site at bp 25,309 and the start of the fiber gene. The primer pair used for the PCR were P-0005/U and P-0010/L. The DNA sequence of P-0005/U was as follows: 5' C TCT AGA AAT GGA CGG AAT TAT TAC AG 3' (SEQ ID No. 17). The sequence corresponds to bp 25,308 through 25,334 of pSQ1 and overlaps the XbaI site in that region. The DNA sequence of P-0010/L was as follows: 5' TTC TTT TCA T CTG CAA CAA CAT GAA GAT AGT G 3' (SEQ ID No. 31). It contains a 5' extension corresponding to the first 10 bp of the Ad41s fiber gene. The remainder of the primer anneals to pSQ1 immediately 5' of the ATG start codon of the fiber gene. The PCR product was the expected size (583 bp). A second PCR was used to amplify the Ad41s fiber using the plasmid pDV60Ad41sF as a template. The primers used were P-0011/U and P-0012/L. The DNA sequence of P-0011/U was as follows: 5' GT TGT TGC AG ATG AAA AGA ACC AGA ATT GAA G 3' (SEQ ID No. 32). It contains a 10 bp 5' extension corresponding to the DNA sequence immediately 5' of the ATG start codon of the fiber gene in pSQ1. The remainder of the primer anneals to the beginning of the Ad41s fiber gene in pDV60Ad41sF. The DNA sequence of P-0012/L was as follows: 5' TG CAA TTG AAA AAT AAA CAC GTT GAA ACA TAA CAC AAA CGA TTC TTT ATT C TTC AGT TAT GTA GCA AAA TAC A 3' (SEQ ID No. 33). It contains a 51 bp

-79-

5' extension corresponding to the sequence in pSQ1 from the last codon of the fiber gene through the MunI site 40 bp downstream of the fiber gene. The remainder of the primer anneals to the 3' end of the Ad41s fiber gene in pDV60Ad41sF. The PCR product was the expected size (1219 bp). The two
5 PCR products were joined by PCR SOEing using primers P-0005/U and P-0009/L. The DNA sequence of P-0009/L was described above. The PCR SOEing reaction yielded the expected 1782 bp product. The product was cloned into pCR4blunt-TOPO to yield pCR4blunt-TOPOAd41sF. Next, pCR4blunt-TOPOAd41sF was digested with XbaI and MunI and the 1773 bp
10 fragment containing the Ad41s fiber gene was gel purified. This fragment was ligated with the 6477 bp XbaI to MunI fragment of pFBshuttle(EcoRI) to generate pFBshuttleAd41sF. The Ad41s fiber gene was transferred into the pSQ1 backbone as follows. First, pFBshuttleAd41sF was linearized using EcoRI and this fragment was ligated with the 24,213 bp EcoRI fragment of pSQ1 to
15 generate pSQ1Ad41sF (Figure 21A). Adenoviral vector containing the Ad41s fiber was generated using the co-transfection strategy described above.

Construction of Ad5 adenoviral vectors containing the Ad41 short fiber with a cRGD targeting ligand in the HI loop: A PCR SOEing strategy was used to generate a construct containing the Ad41s fiber with cRGD in the HI loop. The
20 plasmid pFBshuttleAd41sF was used as a template for the PCR amplifications. First, a 1782 bp fragment was amplified using primers 5FF and 41sRGDR. The primer 5FF was described above. It anneals to pFBshuttleAd41sF at the XbaI site upstream of the fiber gene. The DNA sequence of the primer 41sRGDR was as follows: 5' AGT ACA AAA ACA ATC ACC ACG ACA ATC ACA GTT TAT
25 CTC GTT GTA GAC GAC ACT GA 3' (SEQ ID No. 34). It contains a 30 bp 5' extension that encodes the cRGD targeting ligand. The remainder of the primer anneals to pFBshuttleAd41sF from bp 2878 through 2903. A second PCR amplified a 277bp region of pFBshuttleAd41sF using primers 3FR and 41sRGDF. The primer 3FR was described previously. It anneals to pFBshuttleAd41sF at the
30 MunI site downstream of the fiber gene. The DNA sequence of 41sRGDF was as follows: 5' TGT GAT TGT CGT GGT GAT TGT TTT TGT ACT AGT GGG TAT GCT TTT ACT TTT 3' (SEQ ID No. 35). It contains a 30 bp 5' extension that

-80-

encodes the cRGD targeting ligand and is complementary to the extension on 41sRGDR. The remainder of the primer anneals to pFBshuttleAd41sF from bp 2904 through 2924. The two PCR products were joined by PCR SOEing to generate a 2059 bp fragment using primers 5FF and 3FR. The product was
5 digested with XbaI and MluI and the 1803 bp DNA fragment was gel purified. The fragment was ligated with the 6477 bp fragment resulting from digestion of pFBshuttle(EcoRI) with XbaI and MluI. The resulting plasmid was termed pFBshuttleAd41sRGD. This plasmid was linearized by EcoRI digestion and ligated with the 24,213bp EcoRI fragment of pSQ1 to generate pSQ1Ad41sRGD
10 (Figure 21B).

EXAMPLE 13

***In Vivo* Evaluation Of Ad5 Vectors Containing The Ad41 Short Fiber And Derivatives Thereof**

This example evaluates the *in vivo* biodistribution of adenoviral vectors
15 containing 41sF fibers and derivatives thereof to determine whether vectors containing the these fibers ablate liver transduction due to modified shaft regions. A positive control cohort received Av3nBg (see, Gorziglia *et al.* (1996) *J. Virology* 70:4173-4178) or Ad5. β Gal. Δ F/5F, and a negative control group received HBSS. Ad5. β Gal. Δ F/5F is a derivative of the fiberless vector
20 Ad5. β gal. Δ F (ATCC accession number VR2636) modified to express AD5 fiber (see, *e.g.*, International PCT application No. WO0183729).

The Ad5. β Gal. Δ F vector was pseudotyped with the Ad41sF fiber protein and injected *in vivo*. Cohorts of five C57BL/6 mice received each vector via tail vein injection at a dose of 1×10^{13} particles per kg. The animals were sacrificed
25 approximately 72 hours after vector administration by carbon dioxide asphyxiation. Liver, heart, lung, spleen, and kidney were collected from each animal. The median lobe of the liver was placed in neutral buffered formalin to preserve the sample for β -galactosidase immunohistochemistry. In addition, tissue from each organ was frozen to preserve it for hexon PCR analysis to
30 determine vector content. A separate sample of liver from each mouse was frozen to preserve it for a chemiluminescent β -galactosidase activity assay. β -galactosidase immunohistochemistry, hexon real-time PCR and the

-81-

chemiluminescent β -galactosidase activity assay was carried out as described in example 3.

The results of the hexon DNA analysis showed a dramatic reduction in liver transduction by vectors containing the Ad41sF fiber (Figure 22) with an
5 approximately a 5-fold reduction in liver adenoviral DNA content compared to either control vector.

In the above examples, several novel adenoviral vectors were generated containing various fiber modifications designed to ablate the normal tropism of the vector. See Table 3. Vectors were generated in which the heparan sulfate
10 binding domain in the fiber shaft was replaced by amino acid substitutions. This mutation, termed HSP, was also combined with the KO1 mutation (fiber knob AB loop mutation that ablates CAR binding), and the PD1 mutation (penton mutation that eliminates RGD/integrin interaction). In addition, a vector containing all three mutations (HSP, KO1, PD1) was generated. All vectors containing the HSP
15 mutation, either alone or combined with other capsid modifications, showed dramatically reduced transduction efficiencies on A549 and HeLa cells.

Furthermore, the same vectors showed dramatically reduced transduction of the liver following systemic delivery to mice. As an alternative strategy to ablate the normal tropism of Ad5-based vectors, the Ad5 fiber was replaced by a fiber from
20 a different adenovirus serotype which does not bind CAR and does not contain the heparan binding domain in the shaft. Thus, vectors were generated containing the Ad35 fiber and the Ad41 short fiber. Versions of these two vectors containing a cRGD targeting ligand in the HI loop of the fiber were also produced. Additionally, vectors containing chimeric fibers were generated. A
25 vector containing the Ad35 fiber tail and shaft regions fused to the Ad5 fiber knob domain as well as a vector containing the Ad5 fiber tail and shaft fused to the Ad35 fiber knob domain were constructed. Vectors containing either the entire Ad35 or Ad41 short fiber showed a significant reduction in liver transduction following delivery to mice via the tail vein. The observation of
30 reduced liver transduction using vectors containing either an HSP mutation, the Ad35 fiber, or the Ad41 short fiber indicates the feasibility of detargeting adenoviral vectors *in vivo*. *In vitro* data with the Ad35 fiber or the Ad41 short

-82-

fiber with cRGD (see Example 14) indicate that the virus is completely viable, that is, it is not damaged by the absence of an HSP binding site and is retargetable. Taken together these data suggest that these vectors provide a suitable platform for retargeting strategies.

5

TABLE 3
Description Of Recombinant Adenoviral Vectors Used
To Demonstrate That Shaft Modifications Influence Tropism *In Vivo*
Vector

Vector	Description
10 Av1nBg	An E1 and E3-deleted adenoviral vector encoding a nuclear localizing β -galactosidase
Ad5 Fiber derivatives:	
Av1nBgFKO1	The same as Av1nBg but containing the KO1 AB loop mutation in the fiber gene
Av1nBgPD1	The same as Av1nBg but containing the penton PD1 mutation that deletes the integrin binding, RGD tripeptide
Av1nBgS*	The same as Av1nBg but containing the 4 amino acid substitution in the shaft referred to as S* that modifies the HSP binding motif
15 Av1nBgFKO1S*	The same as Av1nBg but containing the fiber KO1 and S* mutations combined
Av1nBgS*PD1	The same as Av1nBg but containing the fiber S* and penton PD1 mutations combined
Av1nBgFKO1S*PD1	The same as Av1nBg but containing the fiber KO1, S* and penton PD1 mutations combined
Ad35 fiber derivatives:	
Av1nBg35F	The same as Av1nBg but containing the full length Ad35 fiber cDNA
20 Av1nBg5T35H	The same as Av1nBg but containing the 5T35H chimeric fiber
Av1nBg35T5H	The same as Av1nBg but containing the 35T5H chimeric fiber
Av1nBg35FRGD	The same as Av1nBg but containing the full length Ad35 fiber cDNA with a cRGD ligand in the HI loop of the Ad35 fiber
Ad41sF fiber derivatives:	
Av1nBg41sF	The same as Av1nBg but containing the full length Ad41 short fiber cDNA

-83-

Vector	Description
Av1nBg41sFRGD	The same as Av1nBg but containing the full length Ad41 short fiber cDNA with a cRGD ligand in the HI loop of the Ad41 short fiber

EXAMPLE 14***In Vitro* Evaluation Of Adenoviral Vectors Containing The Ad41sF With A cRGD Ligand In The HI Loop**

The transduction efficiencies of adenoviral vectors containing the Ad41sF fiber with the cRGD ligand in the HI loop were evaluated on A549 cells. The transduction efficiencies were compared to that of Av1nBg, an adenoviral vector containing wild type fiber or Av1nBgFKO1RGD, an adenoviral vector containing the KO1 mutation in combination with the cRGD ligand in the HI loop. The day prior to infection, cells were seeded into 24-well plates at a density of approximately 1×10^5 cells per well. Immediately prior to infection, the exact number of cells per well was determined by counting a representative well of cells. Each of the vectors, Av1nBg, Av1nBgFKO1RGD, and Av1nBg41sFRGD were used to transduce A549 cells at a particle to cell ratios of 0 up to 10,000. The cell monolayers were stained with X-gal 24 hours after infection and the percentage of cells expressing β -galactosidase was determined by microscopic observation and counting of cells. Transductions were done in triplicate and three random fields in each well were counted, for a total of nine fields per vector. The results (Figure 23) show that the Av1nBg41sFRGD vector transduced cells to an equivalent level as Av1nBgFKO1RGD at all vector doses examined. Neither FKO1 or Ad41sF can bind CAR. The Ad41sF does not normally interact with CAR and additionally does not contain the HSP binding motif within the shaft domain. These data show that targeting peptides inserted into the loop regions of the fiber knob of KO1 and Ad41sF allows for transduction of target cells via the targeted receptor. Surprisingly, HSP, not CAR and integrins, is the major entry route *in vivo* and ablation of HSP binding permits targeting of adenoviral vectors.

-84-

EXAMPLE 15**Effect of the shaft modification on the biodistribution of adenoviral vectors *in vivo***

The influence of fiber and penton modifications on the *in vivo*

- 5 biodistribution of adenoviral vectors containing fiber head, shaft and penton mutations was examined. Vectors containing the HSP mutation combined with KO1, or PD1, or a combination of all three mutations were evaluated as well as vectors containing the KO1 mutation alone and the PD1 mutation alone. The indicated adenoviral vectors were systemically administered to C57BL6 mice as
- 10 described above. A positive control cohort received Av1nBg and a negative control group received HBSS. Cohorts of five C57BL/6 mice received each vector via tail vein injection at a dose of 1×10^{13} particles per kg. The animals were sacrificed approximately 72 hours after vector administration by carbon dioxide asphyxiation. Liver, heart, lung, spleen, and kidney were collected from
- 15 each animal. Tissue from each organ was frozen to preserve it for real time PCR analysis to determine adenoviral hexon DNA content. A separate sample of liver from each mouse was frozen to preserve it for a chemiluminescent β -galactosidase activity assay. Hexon real-time PCR and the chemiluminescent β -galactosidase activity assay was carried out as described in Example 3.
- 20 The results derived from the liver are described in Example 6 (Figure 14A and B) and also shown in Figure 26 with results presented as percent control of Av1nBg. The effect of the S* shaft modification on the biodistribution of adenovirus to the other organs is shown in Figure 25. The average adenoviral DNA content was determined as adenoviral genomic copies per cell and
- 25 expressed as a percentage of the Av1nBg (+) control value. The average percent control value + standard deviation is shown (n=5 per group) for each tissue examined (Figure 25). Systemic delivery of Ad5 based vectors with wild-type fiber results in a preferential accumulation of vector DNA in the liver with 64 copies per cell with significantly less DNA found in the other organs
- 30 with 1.32 copies per cell found in lung, 2.18 copies per cell in spleen, 0.47 copies per cell found in heart, and 0.72 copies per cell in the kidney. All differences found with PD1, S*, KO1PD1, KO1S*, S*PD1, and KO1S*PD1 were

-85-

significantly different than the Av1nBg (+) control using a unpaired, t-test analysis, P value (0.024. When expressed as a percent of the Av1nBg control values, the influence of each mutation, individually or in combination, becomes apparent. The S* mutation dramatically reduced gene transfer to all four organs, 5 whereas, the KO1 mutation did not. Thus, the importance of the shaft for transduction *in vivo* extends to organs besides the liver. Finally, gene transfer to the lung, heart, and kidney was diminished with PD1 suggesting a role for integrin binding in vector entry in these organs.

EXAMPLE 16

10 Retargeting the S*, shaft modification and the 41sF fiber *in vivo*

Vectors containing the HSP mutation have been shown to effectively detarget adenoviral vectors *in vivo* (see examples 6 and 15). The objective of this study was to evaluate the ability to retarget vectors containing the S* modification or the Ad41sF to tumors *in vivo*. A cRGD peptide was genetically 15 incorporated into the fiber HI loop and evaluated *in vitro* (Examples 8 and 14). These same vectors were then evaluated *in vivo* in tumor-bearing mice. Athymic nu/nu female mice were injected with 8×10^6 A549 cells on the right hind flank. When tumors reached approximately 100mm³ in size, they were randomized into treatment groups. Cohorts of 6 mice received each vector via tail vein 20 injection at a dose of 1×10^{13} particles per kg. The animals were sacrificed approximately 72 hours after vector administration by carbon dioxide asphyxiation. Tumor, liver, heart, lung, spleen, and kidney were collected from each animal. Tissue from each organ was frozen to preserve it for real time PCR analysis to determine adenoviral hexon DNA content. Hexon real-time PCR was 25 carried out as described in example 3. A separate sample of liver from each mouse was frozen to preserve it for a chemiluminescent β -galactosidase activity assay. Hexon real-time PCR and the chemiluminescent β -galactosidase activity assay was carried out as described in example 3.

The adenoviral vector biodistribution to the liver and tumor for each 30 treatment group is shown in Figure 27. Vectors containing the S*, KO1S*, and 41sF fibers effectively detargeted the liver and tumor resulting in a significant reduction in the amount of adenoviral DNA found in each tissue in comparison to

-86-

the Av1nBg control. Vectors containing the cRGD targeting ligand restored transduction of the tumors to levels comparable to that achieved with the untargeted vector.

These data demonstrate successful liver detargeting accompanied with tumor retargeting. The extent of tumor retargeting is related to the affinity and type of ligand that is used. These data demonstrate the successful development of a targeted, systemically deliverable adenoviral vector that will target tumors *in vivo*.

EXAMPLE 17

10 Scale-Up Method For The Propagation Of Detargeted Adenoviral Vectors

The growth and propagation of doubly or triply ablated adenoviral vectors requires novel scale up technologies. These detargeted vectors require alternative cellular entry strategies to allow for the efficient growth and generation of high titer preparations. A strategy for vector growth that is generally applicable to all detargeted adenoviral vectors, that does not require the development of new cell lines, and that also can be used for generating targeted vectors is provided herein.

Three recombinant adenoviral vectors were prepared that contain single mutations in the fiber or penton or both mutations combined into one vector.

20 These vectors are designated Av3nBgFKO1, Av1nBgPD1, and Av1nBgFKO1PD1, respectively. The construction of these vectors is described above and a general description of each vector can be found in Table 1 above.

Scale-up of detargeted adenoviral vectors: A polycation, specifically hexadimethrine bromide was obtained from Sigma Chemical Co (St. Louis, MO),

25 Catalog No. 52495, and was maintained in the medium at 4 μ g/ml during the course of transfections and infections. To illustrate the effects of hexadimethrine bromide on the yield of detargeted adenoviral vectors the following experiment was carried out. Seven plates of AE1-2a adenoviral producer cells (Gorziglia *et al.* (1996) *J. Virology* 70:4173-4178) were
30 transduced with 10 particles per cells of each of the indicated vectors (See Table 4). Each vector was incubated with medium (Richters with 2% HI-FBS) containing hexadimethrine bromide at 4 μ g/ml for 30 min at room temperature

-87-

prior to infection. The infection was carried out for 2 hrs. Complete medium containing hexadimethrine bromide at 4 $\mu\text{g/ml}$ was added to each plate. Final concentration of hexadimethrine bromide in all of these experiments was maintained at 4 $\mu\text{g/ml}$. The titers were determined spectrophotometrically using the conversion of 1OD at A260nm per 1×10^{12} particles (Mittereder *et al.* (1996) *J Virology* 70:7498-7509). The total particle yield was then normalized for the number of plates used for transduction.

The inclusion of hexadimethrine bromide in the medium during the course of infection allows for the efficient propagation of detargeted adenoviral vectors containing fiber and penton mutations either alone or in combination. The affect of hexadimethrine bromide on vector yields is shown in Table 4. A 35-fold improvement in the yield of Av3nBgFKO1 was found when hexadimethrine bromide was included in the culture medium and resulted in increasing the vector yield from 1.3×10^{10} up to 4.6×10^{11} vector particle per plate. Hexadimethrine bromide has a minimal effect on the yield of the Av1nBgPD1 adenoviral vector containing the penton, PD1 mutation with only a 1.2 fold improvement. The greatest effect using hexadimethrine bromide was found on the propagation of the doubly ablated adenoviral vector, Av1nBgFKO1PD1 with increases in vector yield from barely detectable levels up to 4.53×10^{10} vector particles per plate. These data demonstrate that use of nonspecific entry mechanisms allows for the efficient scale-up of detargeted adenoviral vectors.

TABLE 4
Efficient Scale-Up Of Detargeted Adenoviral Vectors Using hexadimethrine bromide

Vector	Vector Yield (particles/plate)		Fold Improvement
	(-) hexadimethrine bromide	(+) hexadimethrine bromide	
Av1nBg	3.89×10^{11}	5.72×10^{11}	1.47
Av3nBg	8.58×10^{10}	2.38×10^{11}	2.77
Av3nBgFKO1	1.30×10^{10}	4.60×10^{11}	35.4
Av1nBgPD1	1.95×10^{11}	2.40×10^{11}	1.23
Av1nBgFKO1PD1	TLTC*	4.53×10^{10}	1

-88-

*TLTC: Too low to count, a faint virus band was collected and the particle concentration was too dilute for titer determination.

† Significant improvement

The use of alternative polycations including protamine sulfate and poly-lysine as well as bifunctional proteins such as the anti-penton:TNF α fusion protein was investigated. Figure 24 show results that demonstrate all the reagents tested had some effect on enhancing transduction of the Av3nBgFKO1 vector. All of these compounds, when maintained in the medium during infection, enhanced transduction of the Av3nBgFKO1 detargeted adenoviral vector.

Bifunctional reagents: The use of bifunctional reagents for the propagation of detargeted adenoviral vectors was examined using the anti-penton:TNF α fusion protein. This particular reagent is a fusion protein between an antibody against Ad5 penton and the TNF α protein that is produced using stably transfected insect cells. This reagent will bind specifically to the adenoviral capsid via penton base and allow for binding to cell surface TNF receptors. The use of this reagent for the propagation of detargeted vectors is illustrated in Table 5 using Av3nBgFKO1 (also shown in Figure 24). Monolayers of S8 cells were infected with 10 or 100 particles per cell of Av3nBgFKO1 or a control vector in the presence or absence of 1 μ g/ml of the anti-penton:TNF α fusion protein. The monolayers were visually inspected over time for vector spread as indicated by the extent of cytopathic effect (CPE). The percentage of CPE at each time point is shown. The use of this bifunctional reagent clearly enhances the spread of the Av3nBgFKO1 vector throughout the monolayer.

TABLE 5
Efficient Scale-Up Of Detargeted Adenoviral
Vectors Using Bifunctional Reagents: Anti-Penton:TNF α

	10 ppc - anti-penton TNF	10 ppc + anti-penton TNF	100 ppc - anti-penton TNF	100 ppc + anti-penton TNF
	Percentage of CPE			
Ad5Luc1				
24 h	0%	0%	0%	0%

-89-

	10 ppc - anti-penton TNF	10 ppc + anti-penton TNF	100 ppc - anti-penton TNF	100 ppc + anti-penton TNF
	Percentage of CPE			
48 h	20-30%	20-30%	90-100%	90-100%
72 h	60-70%	80-90%	100%	100%
120 h	100%	100%	100%	100%
Av3nBgKO1 24hrs				
24 h	0%	0%	0%	0%
48 h	0%	10-20%	0%	90-100%
72 h	5%	60-70%	5%	100%
120 h	40-50%	100%	100%	100%

Since modifications will be apparent to those of skill in this art, it is intended that this invention be limited only by the scope of the appended claims.

-90-

WHAT IS CLAIMED IS:

1. A modified adenovirus capsid protein,
the unmodified capsid protein binds to heparin sulfate
proteoglycan (HSP); and
- 5 the capsid protein comprises a mutation, whereby binding to
heparin sulfate proteoglycan (HSP) is altered.
2. The modified protein of claim 1 that is a fiber protein
3. The capsid protein of claim 2, wherein the binding of the modified
fiber protein is eliminated or reduced compared to the unmodified protein.
- 10 4. The modified protein of claim 2, wherein the binding of the
modified fiber protein is eliminated or reduced compared to the unmodified
protein.
5. The modified protein of claim 3 that comprises an insertion,
deletion or replacement of amino acids.
- 15 6. The modified protein of claim 2, wherein the mutation alters the
motif that binds to HSP, whereby HSP interaction is altered.
7. The modified protein of claim 6, motif is BBXB or BBBXB,
wherein the B is a basic amino acid and X is any amino acid.
8. The modified protein of claim 7, wherein the motif comprises the
20 consensus sequence KGTK.
9. The modified protein of claim 2, wherein the fiber is a modified
Ad5 or Ad2 fiber.
10. A modified protein of claim 2 that is a chimeric fiber protein,
comprising portions of fiber proteins from at least two different adenoviruses,
25 wherein:
a shaft or portion thereof is from a first adenovirus, whereby the resulting
fiber does not bind to HSP or binds to HSP with reduced affinity compared to an
unmodified fiber protein;
a shaft or portion thereof from the first adenovirus does not bind
30 to HSP or binds to HSP with reduced affinity compared to the second
adenovirus;
the second adenovirus binds to HSP; and

-91-

the portion comprises a sufficient portion to alter HSP binding of the resulting protein.

11. The modified protein of claim 10, wherein the binding to HSP of the modified fiber protein is eliminated or reduced compared to the unmodified protein.

12. The modified protein of claim 10, wherein the remainder of the fiber protein is from the second adenovirus.

13. The modified protein of any of claims 2, 3, 10 and 11, further comprising one or more further modifications that reduce or eliminate interaction of the resulting fiber with one or more cell surface proteins in addition to HSP.

14. The modified protein of claim 13, further comprising a ligand, whereby the resulting fiber binds to a receptor for the ligand.

15. The modified protein of claim 14, wherein the ligand is included in the knob region.

16. The modified protein of claim 14, wherein the ligand is inserted or it replaces a portion of the fiber, whereby the resulting fiber binds to a receptor for the ligand.

17. A modified protein of claim 11, wherein affinity for HSP is reduced at least by an amount selected from among reduced 5-fold, 10-fold and 100-fold.

18. The modified protein of claim 11, wherein the first adenovirus is selected from the group consisting of subgroup B, D or F, and the second is of subgroup C.

19. The modified protein of claim 10, wherein the first adenovirus is selected from the group consisting of Ad3, Ad35, Ad7, Ad11, Ad16, Ad21, Ad34, Ad40, Ad41 and Ad46.

20. The modified protein of claim 18, wherein the second adenovirus is Ad5 or Ad2.

21. The modified protein of claim 19, wherein the second adenovirus is Ad5 or Ad2.

22. A modified protein of claim 1 selected from the group consisting of a fiber protein comprising:

-92-

the sequence of amino acids set forth in any of SEQ ID Nos. 52, 54, 56, 58, 62, 66, 70 and 72; or

a sequence of amino acids having 90% sequence identity with a sequence of amino acids set forth in any of SEQ ID Nos. 52, 54, 56, 58, 62, 66,

5 70 and 72; or

a sequence of amino acids encoded by a sequence of nucleotides that hybridizes under conditions of high stringency along at least 70% of its length to a sequence of nucleotides that encodes a sequence of amino acids set forth in any of SEQ ID Nos. 52, 54, 56, 58, 62, 66, 70 and 72.

10 23. A nucleic acid molecule encoding a modified protein of any of claims 1-12 and 14-22.

24. A nucleic acid molecule encoding a modified protein of claim 13.

25. The nucleic acid molecule of claim 23 that comprises a vector.

26. The nucleic acid molecule of claim 24 that comprises vector.

15 27. The nucleic acid molecule of claim 25 that is an adenovirus vector.

28. The nucleic acid molecule of claim 26 that is an adenovirus vector.

29. The vector of claim 27 that is an adenoviral vector from a subgroup B, C or D adenovirus.

20 30. The vector of claim 28 that is an adenoviral vector from a subgroup B, C or D adenovirus.

31. A cell, comprising a nucleic acid molecule of claim 23.

32. A cell, comprising a nucleic acid molecule of claim 24.

33. The cell of claim 31 that is a eukaryotic cell.

34. The cell of claim 32 that is a eukaryotic cell.

25 35. A cell, comprising a nucleic acid molecule of claim 27, wherein:
the cell is a eukaryotic cell; and
the cell in a packaging cell.

30 36. A cell, comprising a nucleic acid molecule of claim 28, wherein:
the cell is a eukaryotic cell; and
the cell in a packaging cell.

-93-

37. An adenoviral particle, comprising a modified protein of any of claims 1-12 and 14-22, whereby binding of the viral particle to HSP is altered compared to a particle that expresses an unmodified fiber.

5 38. An adenoviral particle, comprising a modified protein of claim 13, whereby binding of the viral particle to HSP is altered compared to a particle that expresses an unmodified fiber.

39. An adenoviral particle of claim 37, wherein a native receptor for the fiber is coxsackie-adenovirus receptor (CAR).

10 40. The adenoviral particle of claim 39, further comprising a mutation in the CAR-binding region of the capsid.

41. The adenoviral particle of claim 39, further comprising a mutation in the α_v integrin-binding region of the capsid, whereby binding to the integrin is eliminated or reduced.

15 42. The adenoviral particle of claim 40, further comprising a mutation in the α_v integrin-binding region of the capsid, whereby binding to the integrin is eliminated or reduced

43. The adenoviral particle of claim 39, wherein the CAR-binding region of the capsid modified is on a fiber knob.

20 44. The adenoviral particle of claim 43, wherein the fiber knob modification is in the AB loop or CD loop.

45. The adenoviral particle of claim 44, wherein the fiber knob modification is selected from the group consisting of KO1 and KO12.

46. The adenoviral particle of claim 39, wherein the adenovirus is a subgroup C, D or F adenovirus.

25 47. The adenoviral particle of claim 46, wherein the subgroup C virus is Ad2 or Ad5, the subgroup D virus is Ad46 and the subgroup F virus is Ad41.

48. The adenoviral vector of claim 27 that is an early generation adenoviral vector, a gutless adenoviral vector or a replication-conditional adenoviral vector.

30 49. The adenoviral vector of claim 28 that is an early generation adenoviral vector, a gutless adenoviral vector or a replication-conditional adenoviral vector.

-94-

50. The adenoviral vector of claim 48, wherein the replication-conditional adenoviral vector is an oncolytic adenoviral vector.

51. The adenoviral vector of claim 49, wherein the replication-conditional adenoviral vector is an oncolytic adenoviral vector.

5 52. The adenoviral vector of claim 27 that comprises heterologous nucleic acid.

53. The adenoviral vector of claim 28 that comprises heterologous nucleic acid.

10 54. The adenoviral vector of claim 52, wherein the heterologous nucleic acid encodes a polypeptide.

55. The adenoviral vector of claim 53, wherein the heterologous nucleic acid encodes a polypeptide.

56. The adenoviral vector of claim 52, wherein the heterologous nucleic acid comprises or encodes a regulatory nucleic acid.

15 57. The adenoviral vector of claim 53, wherein the heterologous nucleic acid comprises or encodes a regulatory nucleic acid.

58. The adenoviral vector of claim 52, wherein the heterologous nucleic acid comprises or encodes a promoter or RNA.

20 59. The adenoviral vector of claim 53, wherein the heterologous nucleic acid comprises or encodes a promoter or RNA.

60. The adenoviral vector of claim 59, wherein the promoter is a cell or tissue specific promoter.

61. The adenoviral vector of claim 59, wherein the promoter is operably linked to a gene of an adenovirus essential for replication.

25 62. The adenoviral vector of claim 60, wherein the tissue specific promoter is a tumor specific promoter.

63. The adenoviral vector of claim 58, wherein the polypeptide is a therapeutic polypeptide.

30 64. A method of expressing heterologous nucleic acid in a cell, comprising transducing the cell with an adenoviral vector of claim 57.

65. The method of claim 64, wherein:
the cell is a tumor cell;

-95-

the adenoviral vector is an oncolytic vector; and
the cell is killed .

66. The method of claim 64, wherein the cell is a mammalian cell.

67. The method of claim 64, wherein the cell is a primate cell.

5 68. The method of claim 67, wherein the cell is a human cell.

69. A method of reducing transduction of liver cells by an adenoviral particle, comprising reducing or eliminating binding of the particle to heparin sulfate proteoglycans (HSPs) on the liver cells.

10 70. A scale up method for the propagation of a detargeted adenoviral particle, comprising:

infected a cell capable of replicating, maturing and packaging an adenoviral vector with a detargeted adenoviral vector in the presence of a reagent that results in entry of the adenoviral particle into the cell;

15 culturing the infected cell under conditions suitable for growth, spread and propagation of the adenoviral vector; and
recovering the resulting adenoviral particles.

71. The method of claim 70, wherein the reagent is a polycation.

20 72. The method of claim 71, wherein the polycation is selected from the group consisting of hexadimethrine bromide, polyethylenimine, protamine sulfate and poly-L-lysine.

73. The method of claim 70, wherein the reagent is a bifunctional protein that binds to the adenoviral particle and to a receptor on the cell.

74. The method of claim 73, wherein:
the bifunctional protein is selected from the group consisting of an
25 anti-fiber antibody ligand fusion, an anti-fiber-Fab-FGF conjugate, an anti-penton-antibody ligand fusion, an anti-hexon antibody ligand fusion and a polylysine-peptide fusion, wherein the ligand is a ligand that binds to the receptor.

30 75. The method of any one of claims 70-74, wherein the detargeted adenoviral particle expresses a modified capsid, whereby binding to at least one host cell receptor is reduced or eliminated compared with a wild-type adenovirus.

-96-

76. The method of claim 75, wherein the adenoviral particle is modified to eliminate or reduce binding with one host cell receptor.

77. The method of claim 75, wherein the adenoviral particle is modified to eliminate or reduce binding with two host cell receptors.

5 78. The method of claim 75, wherein the adenoviral particle is modified to eliminate or reduce binding with three host cell receptors.

79. The, method of claim 75, wherein the particle is modified with one or more mutations selected from the group consisting of mutations that reduce or eliminate interactions with one or more of α_v integrins, coxsackie-adenovirus
10 receptors (CAR). and heparin sulfate proteoglycans (HSP).

80. The method of claim 79, wherein the mutation is selected from the group consisting of PD1, KO1, KO12 and S*.

81. The modified protein of claim 2, wherein the mutation is in the shaft of a fiber.

15 82. A modified protein of claim 3, wherein affinity for HSP is reduced at least by an amount selected from among reduced 5-fold, 10-fold and 100-fold.

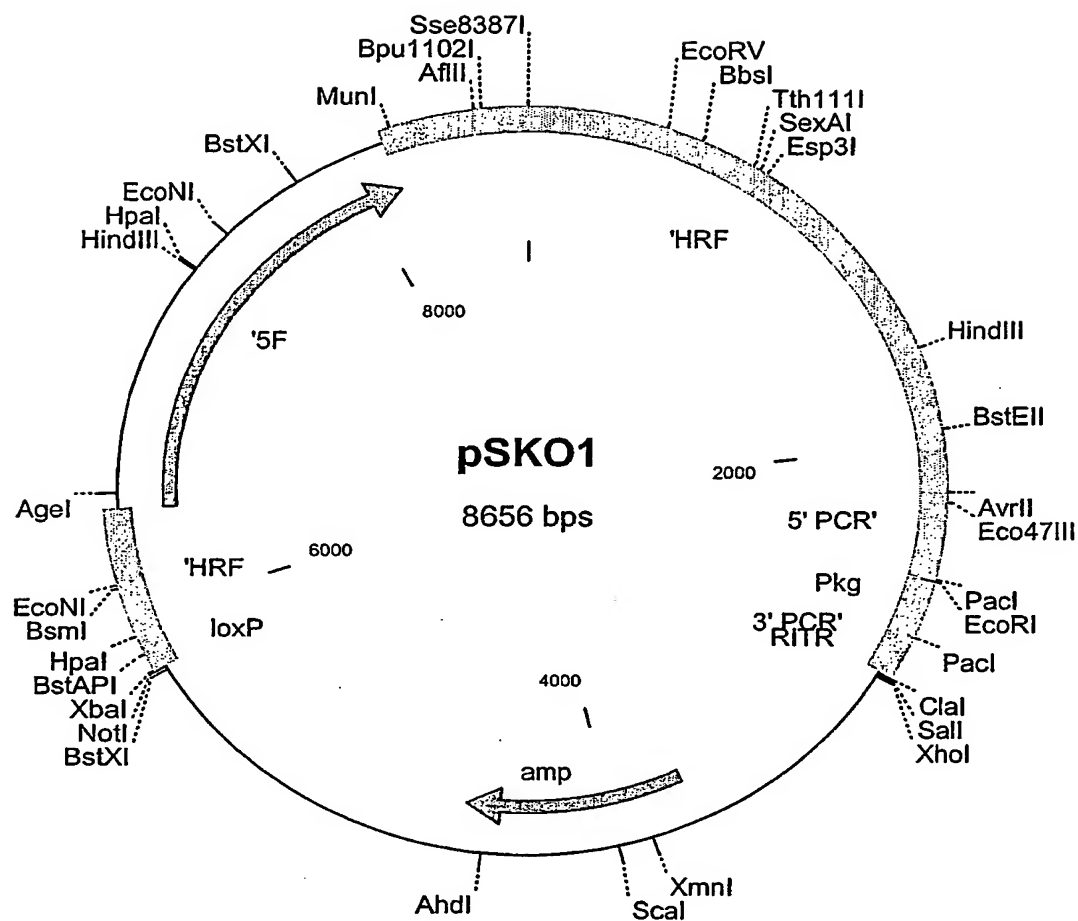
Figure 1

Figure 2

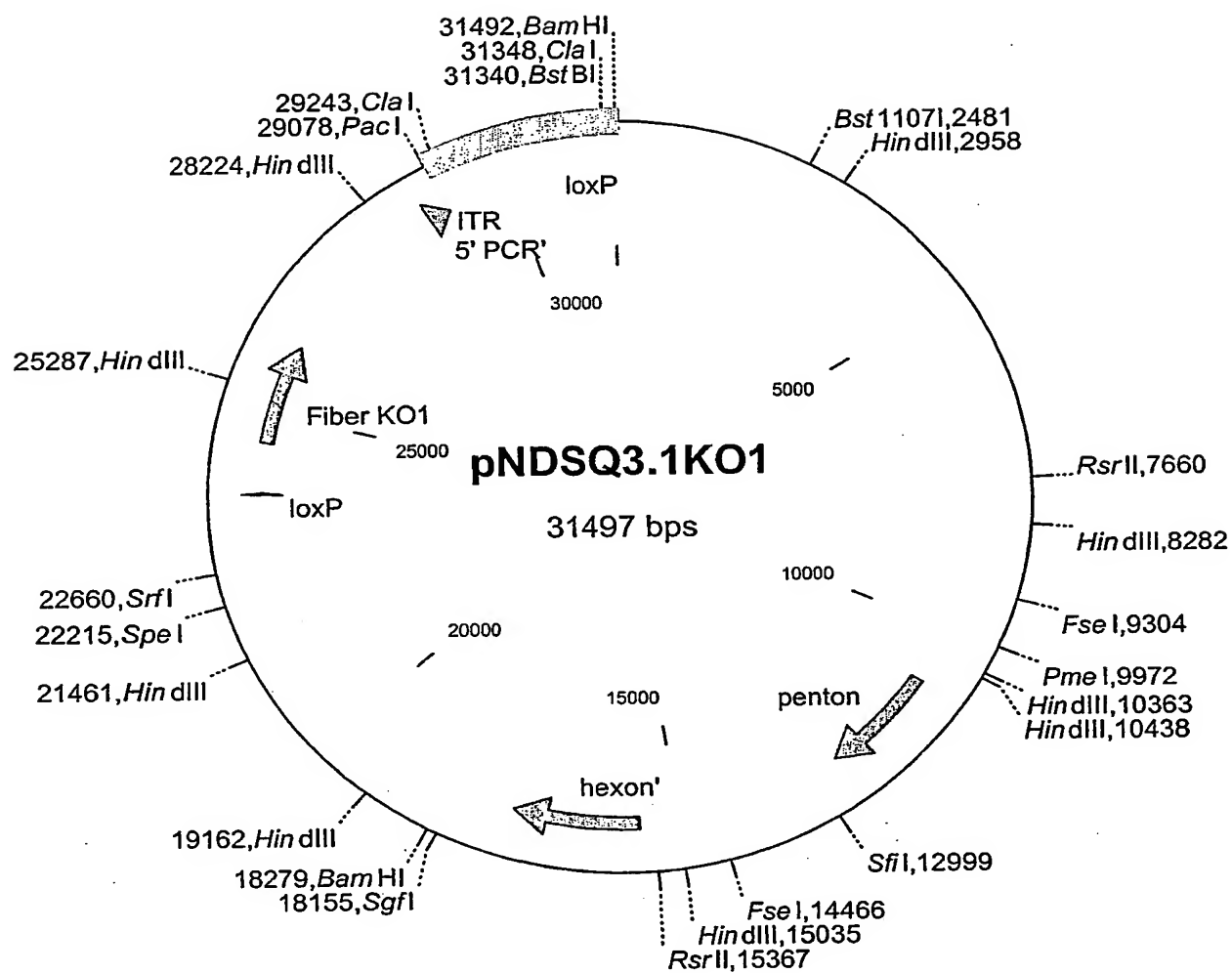


Figure 3A

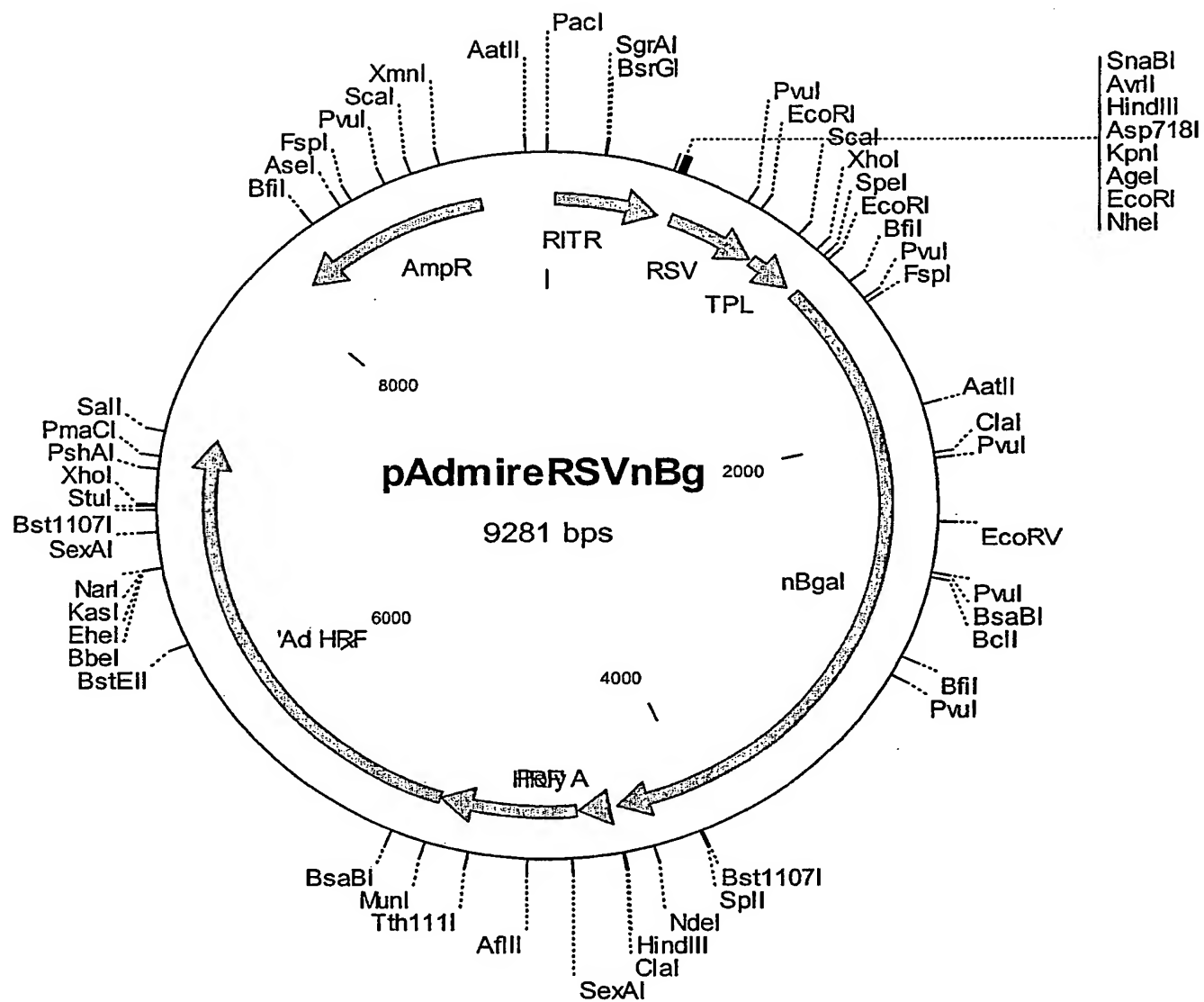


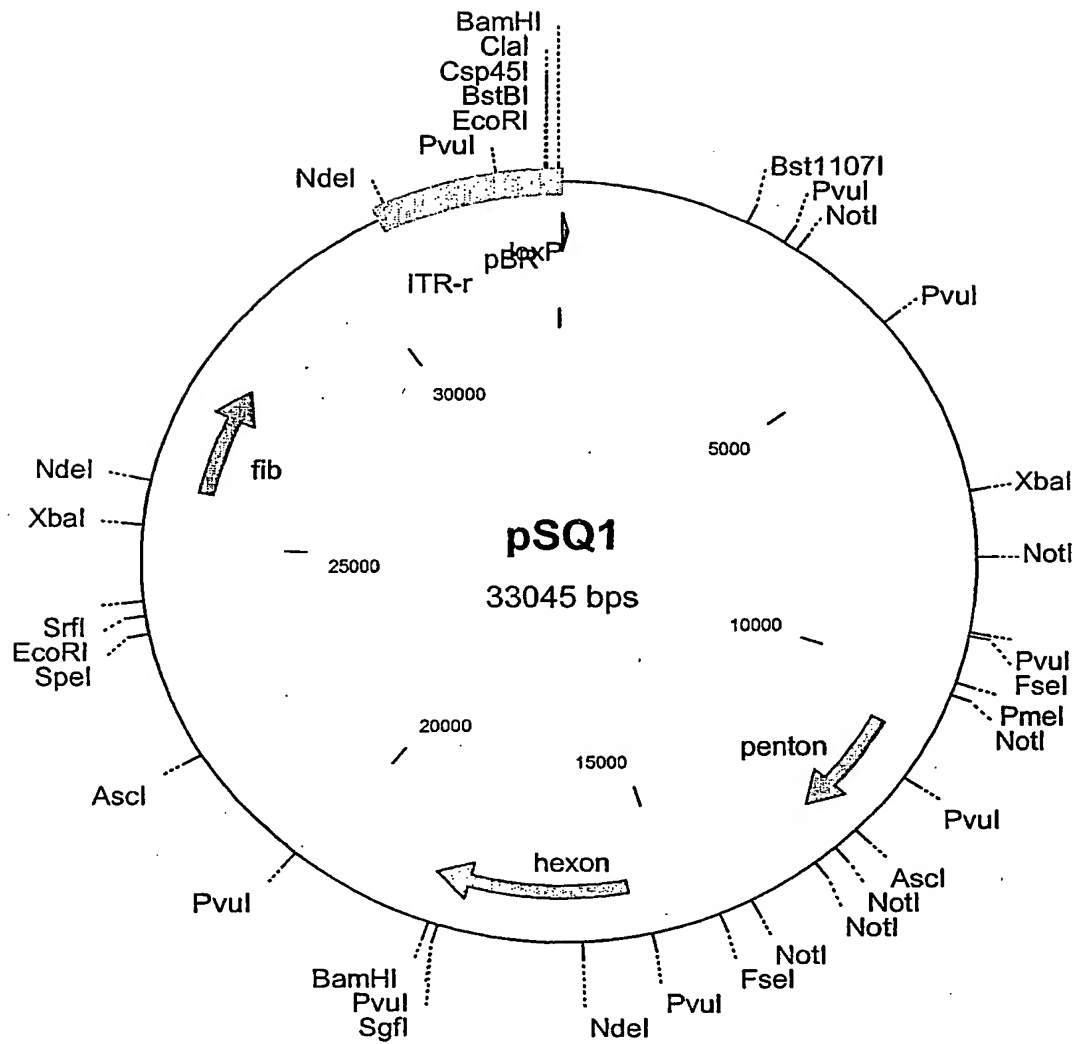
Figure 3B

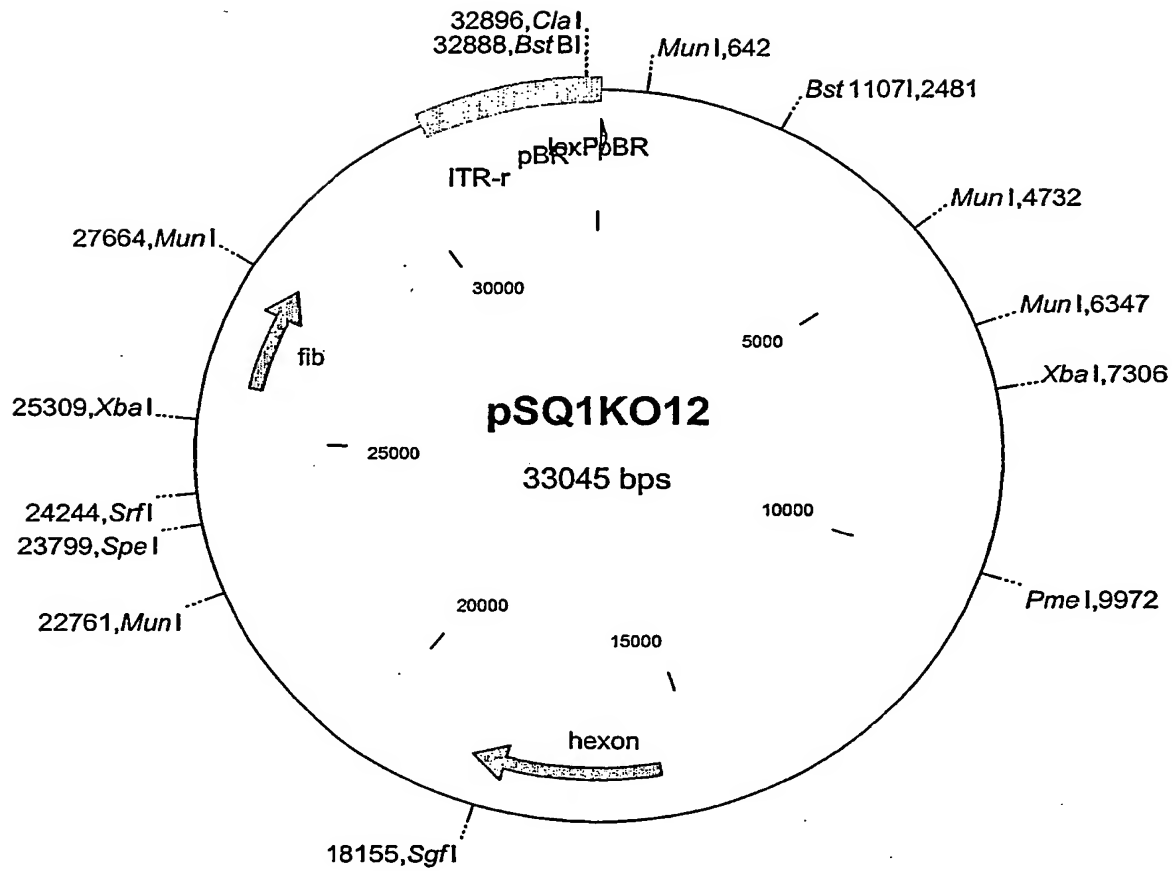
Figure 3C

Figure 4

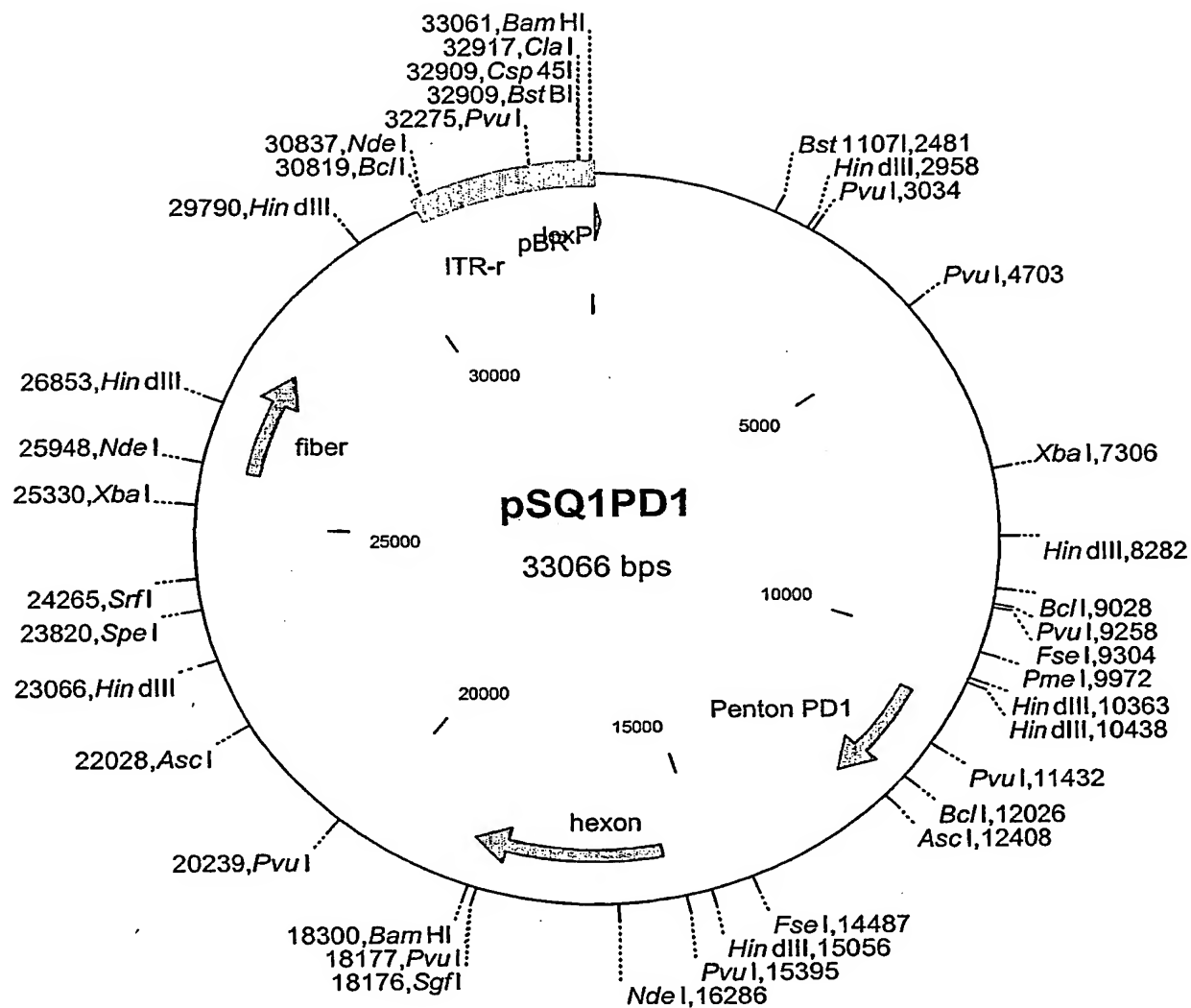


Figure 5A

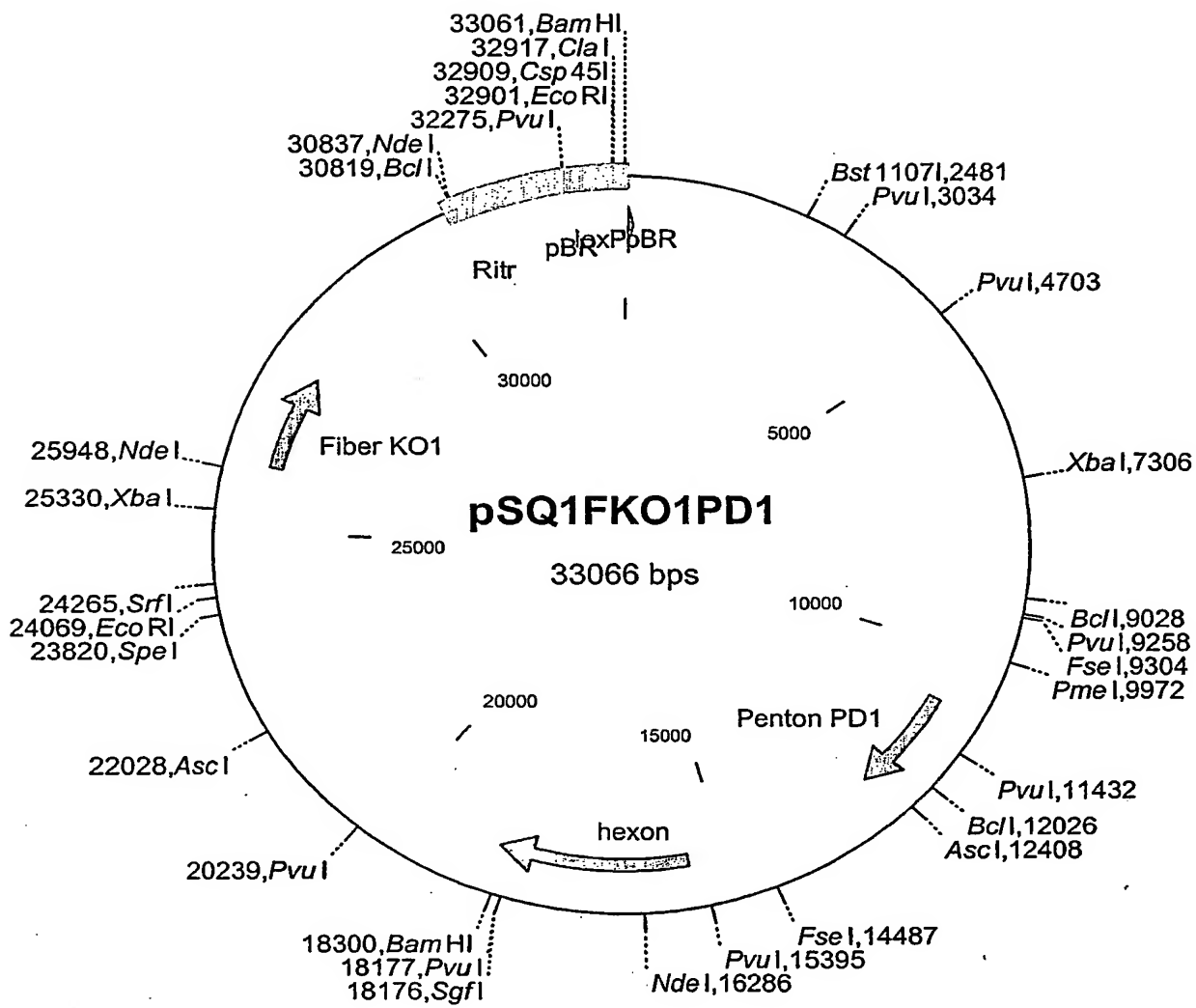


Figure 5B

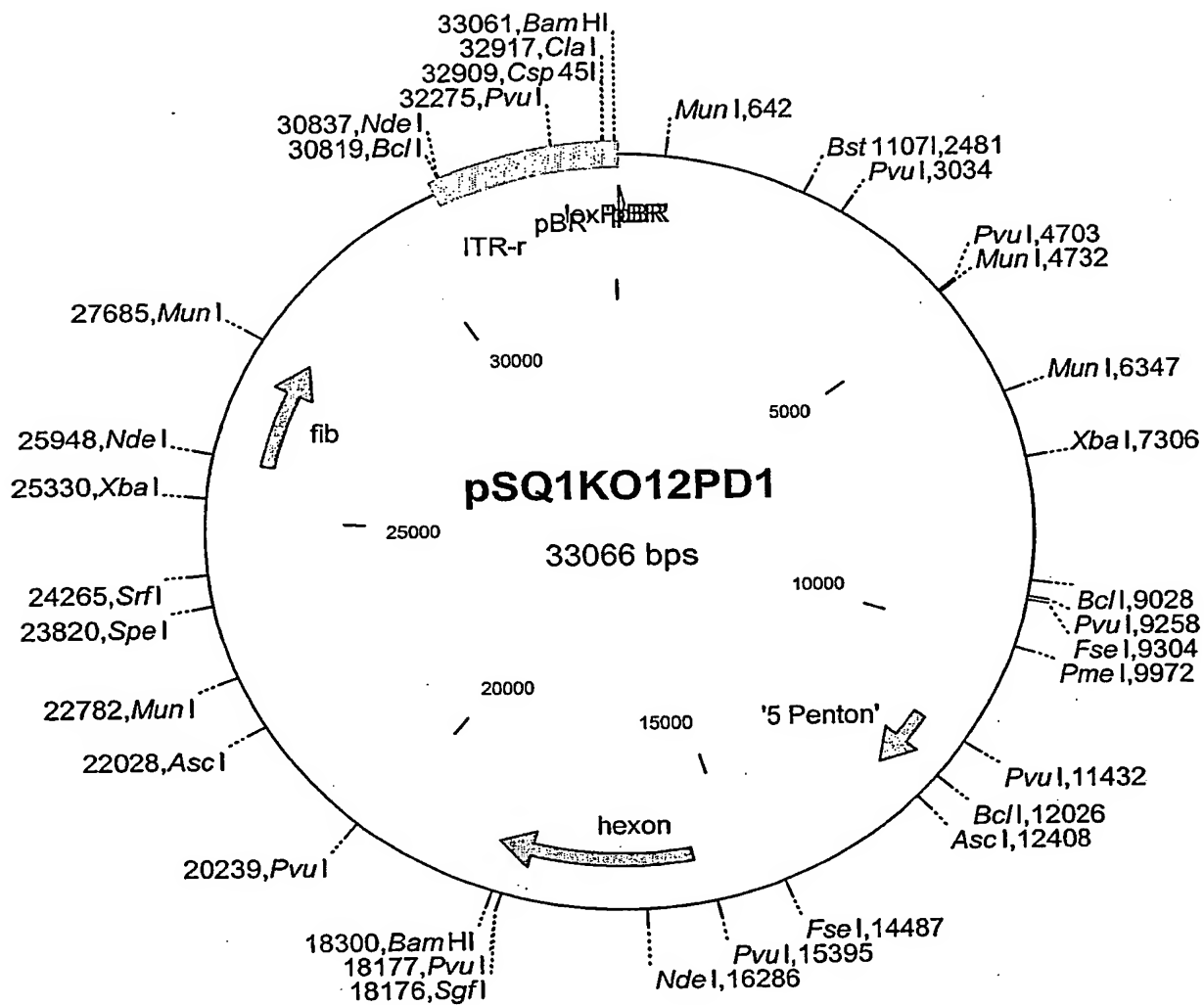
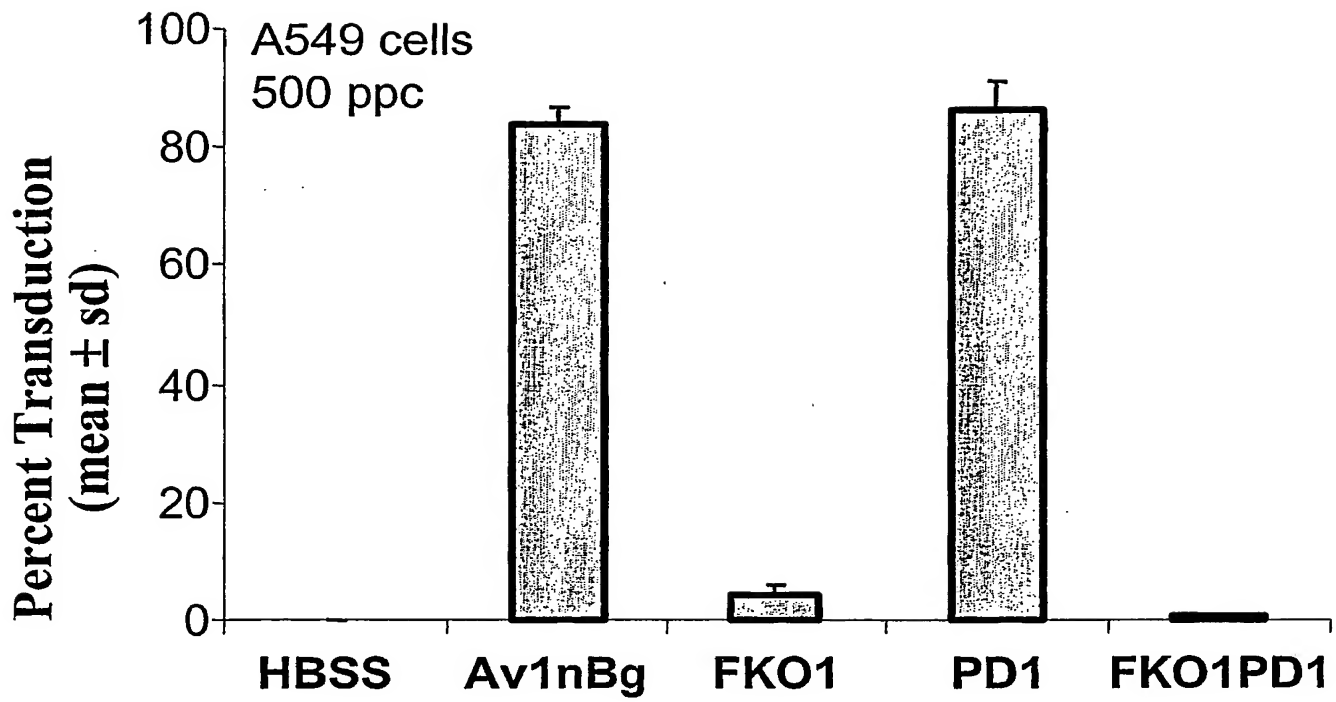


Figure 6



10/35

Figure 7A

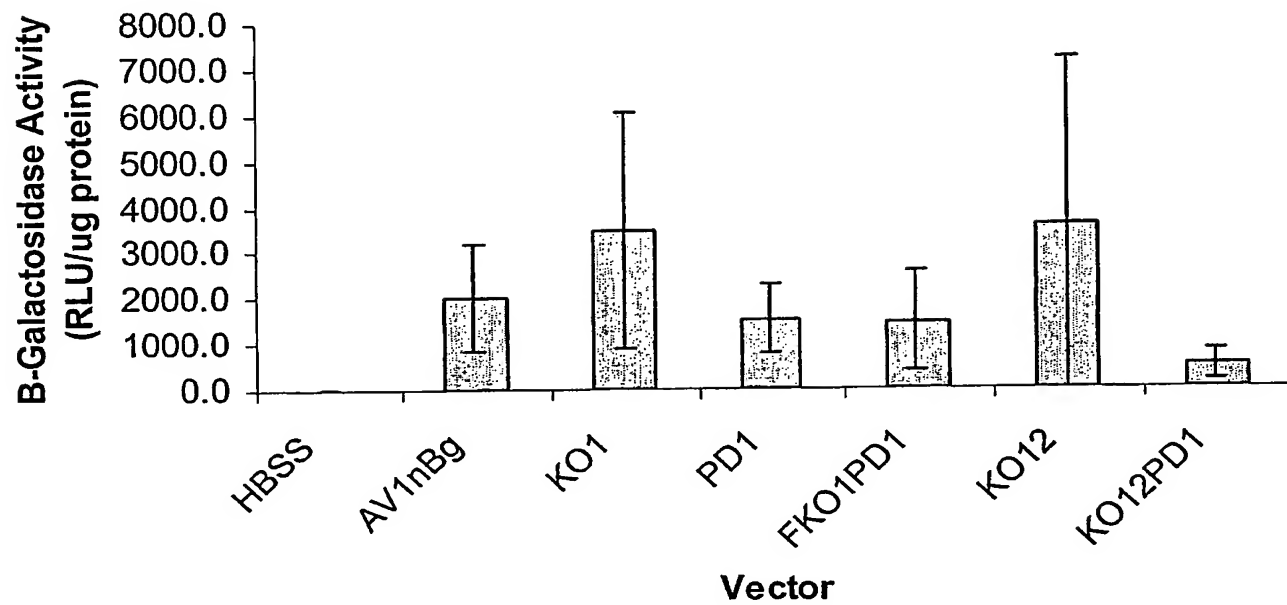


Figure 7B

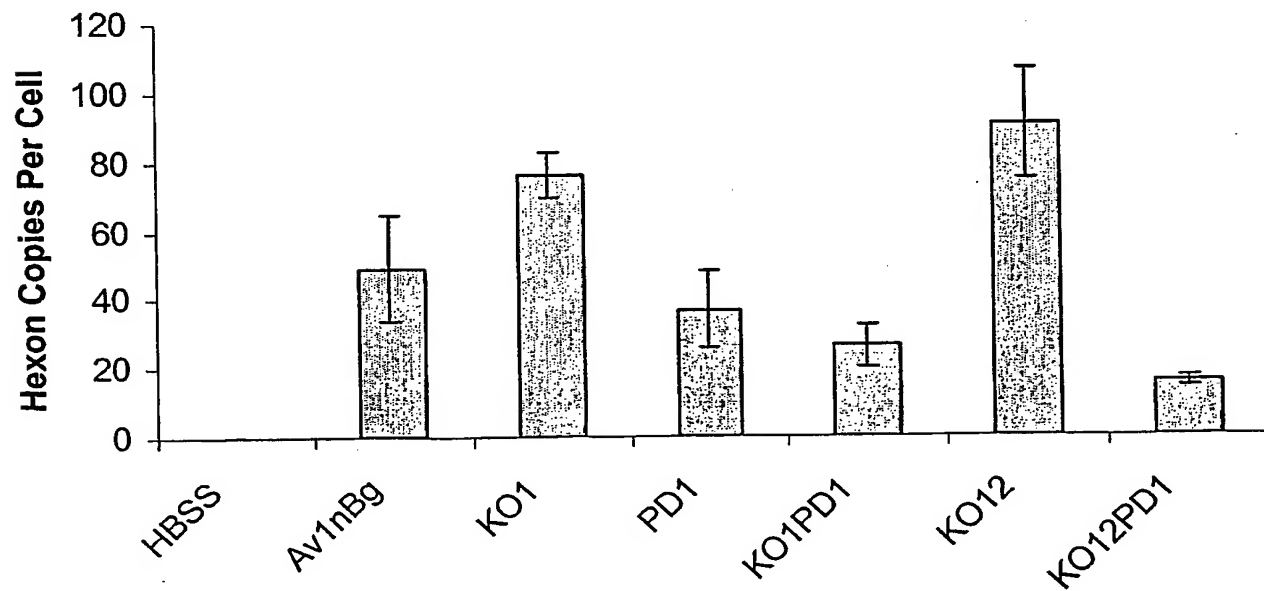


Figure 8

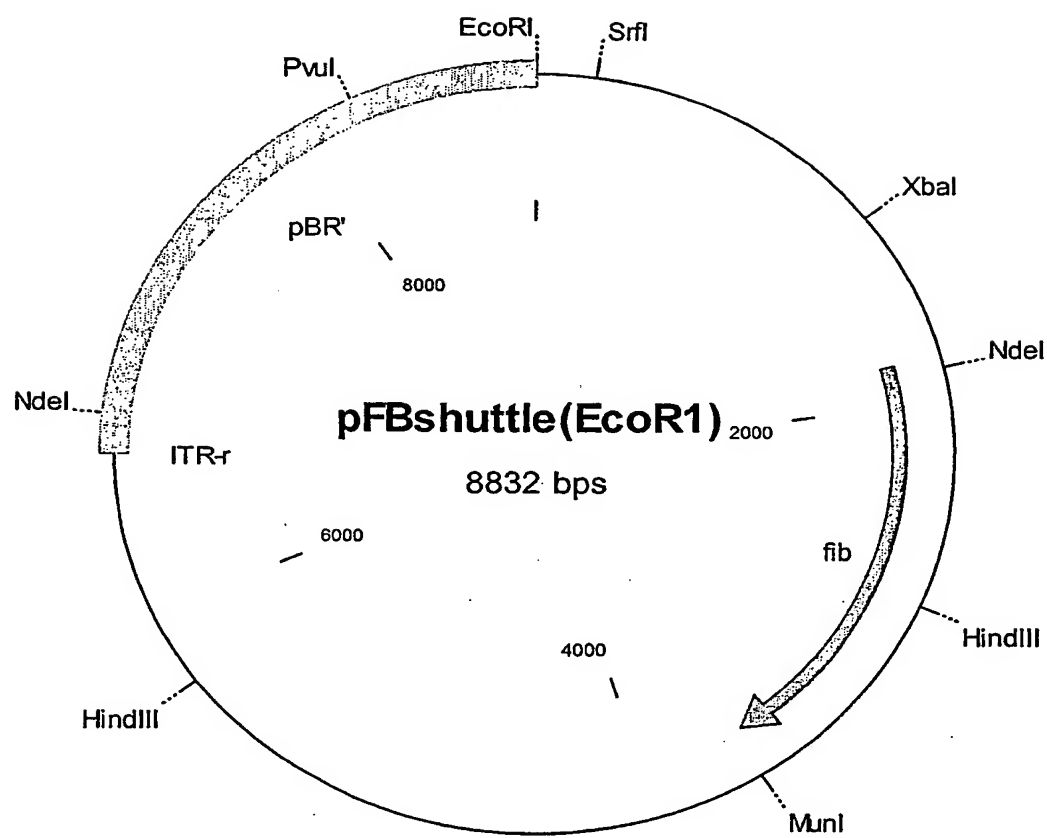


Figure 9

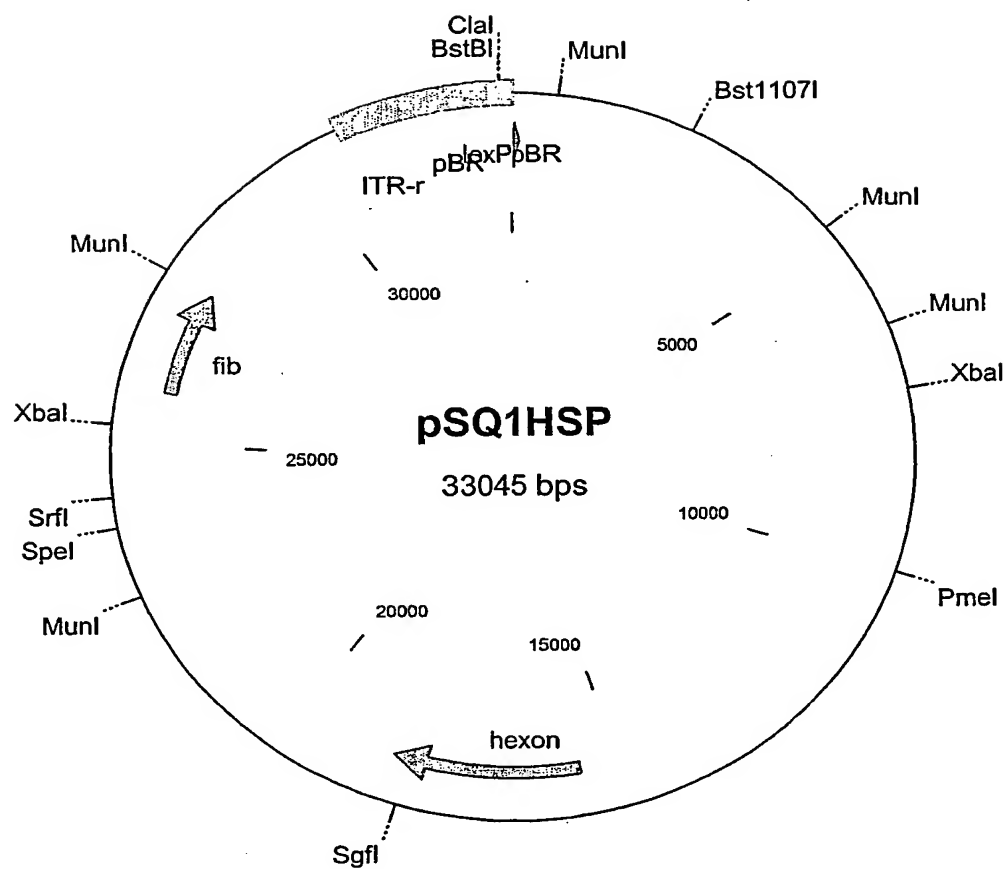


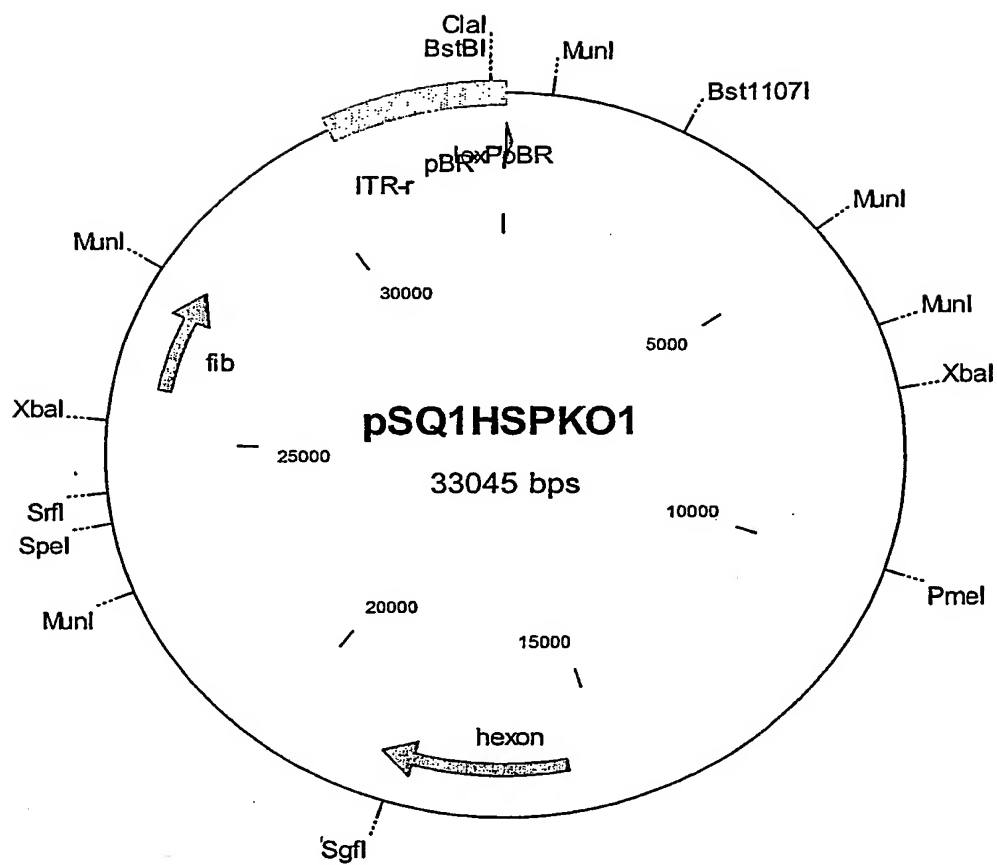
Figure 10

Figure 11

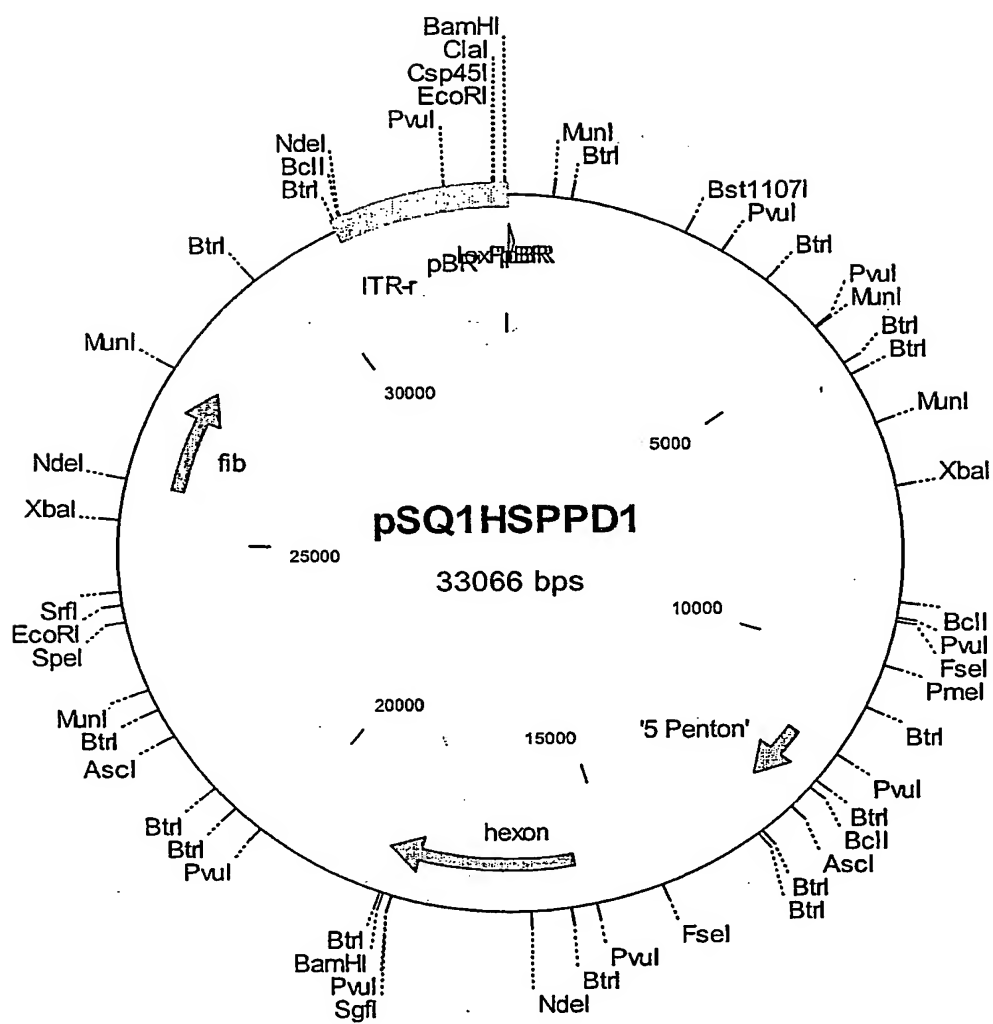


Figure 12

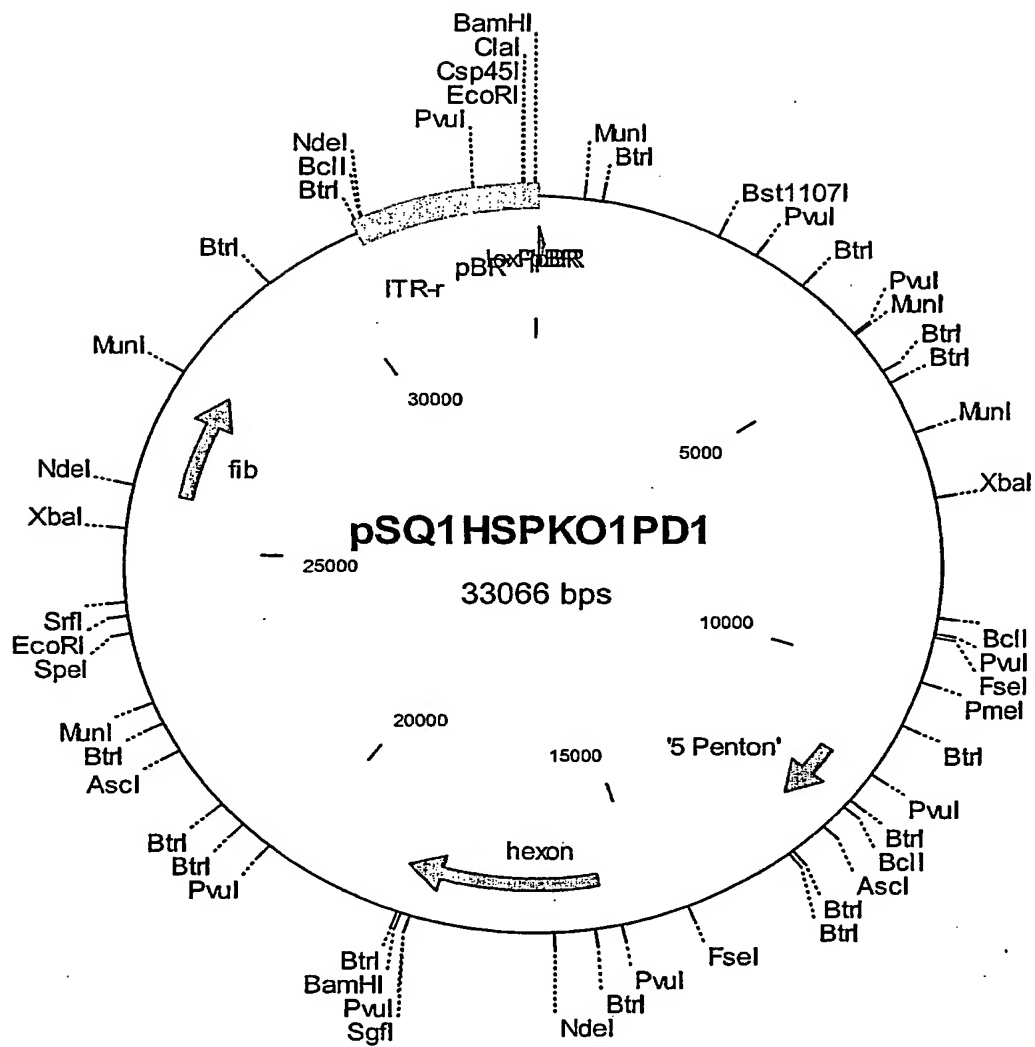


Figure 13A

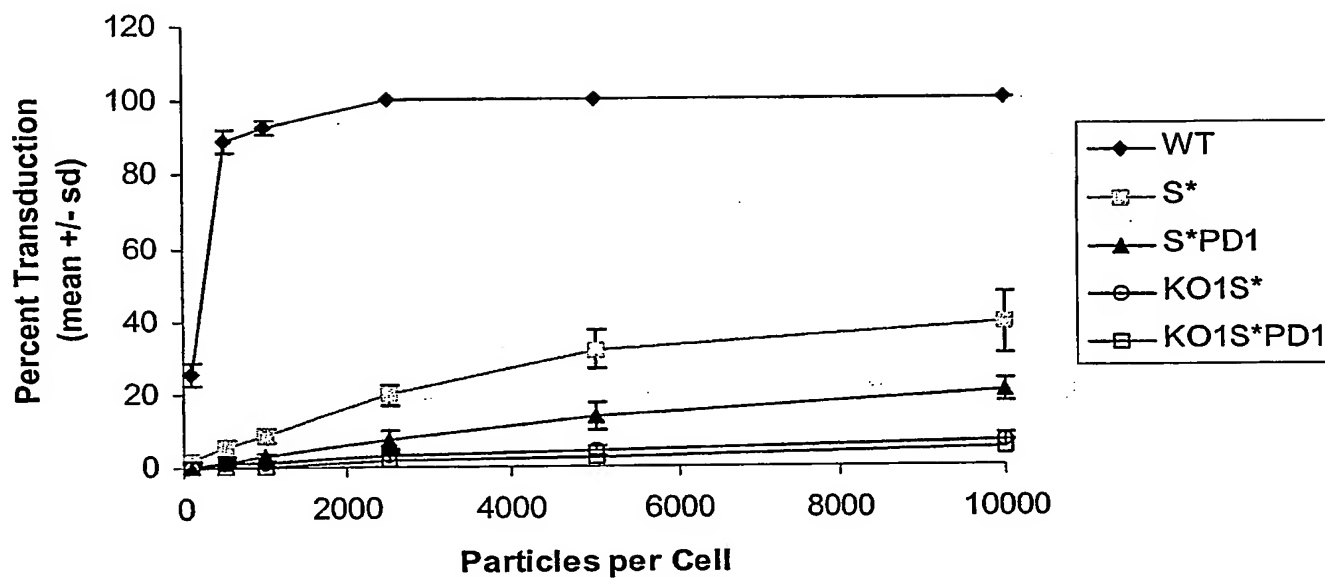


Figure 13B

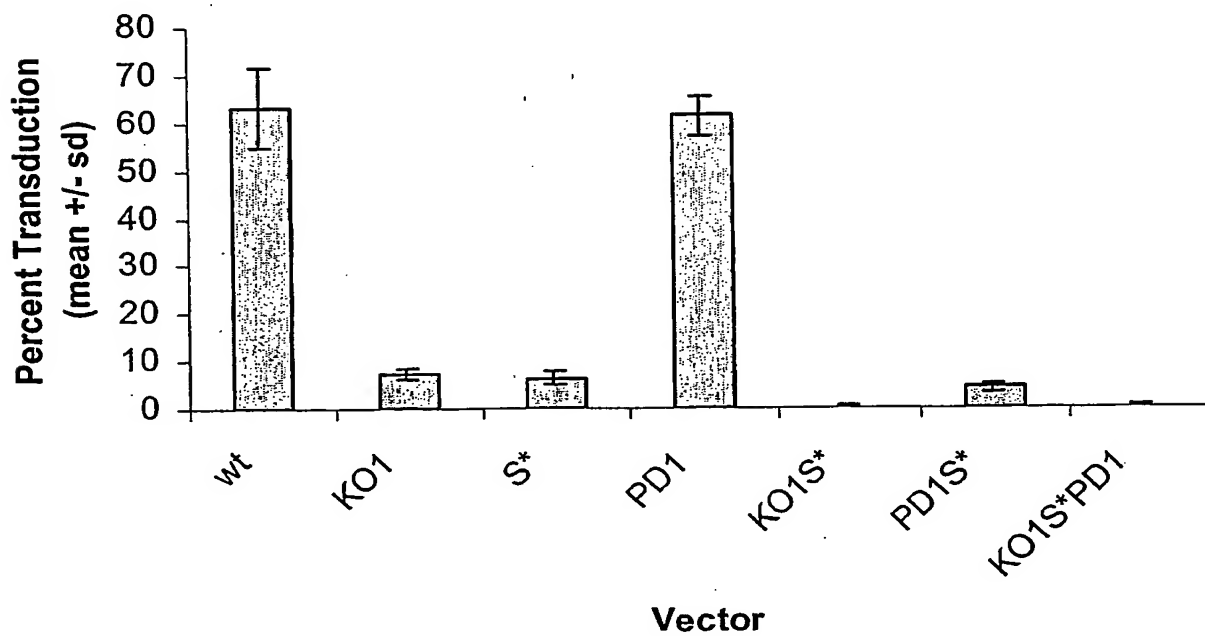


Figure 13C

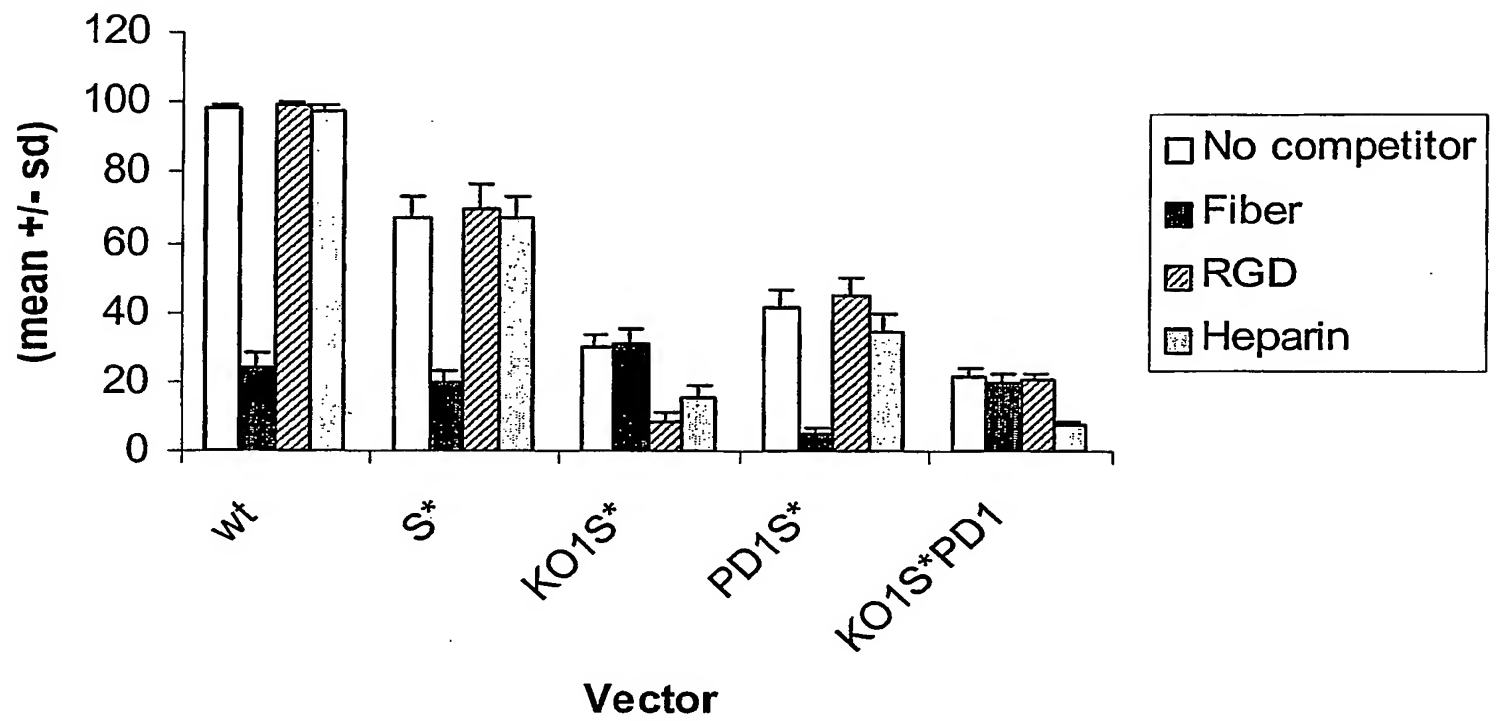


Figure 14A

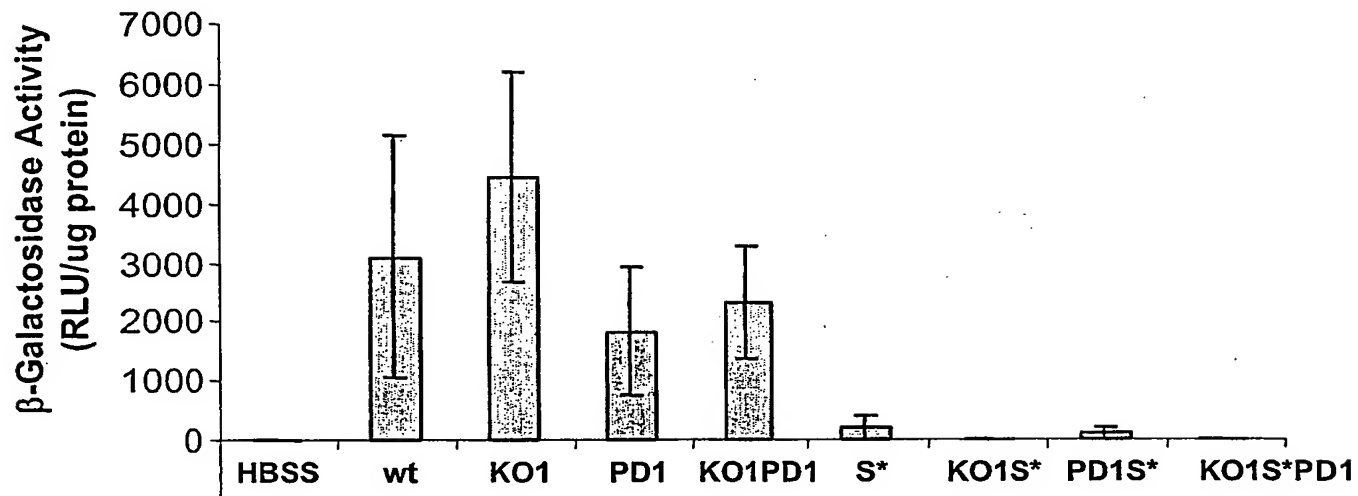


Figure 14B

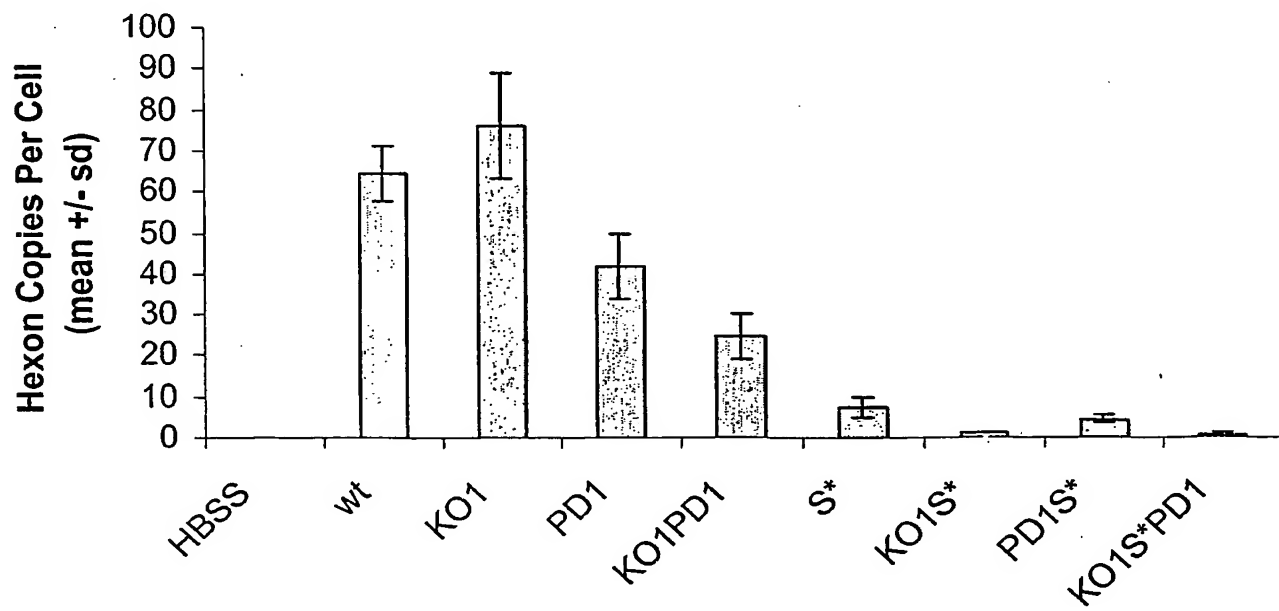


Figure 15A

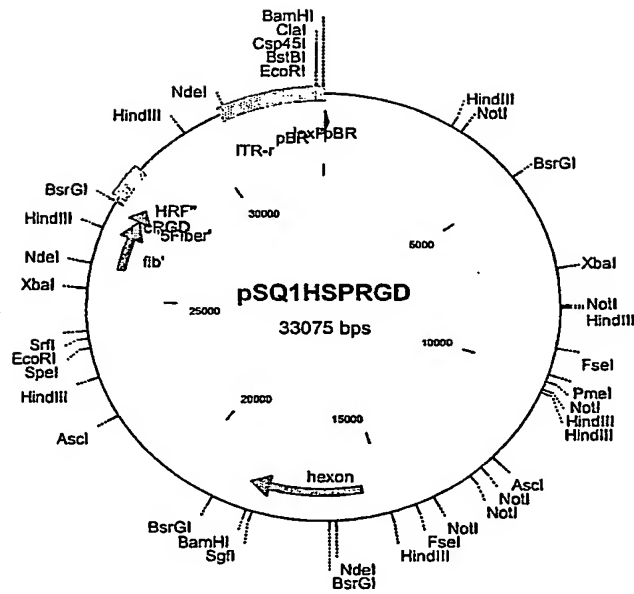


Figure 15B

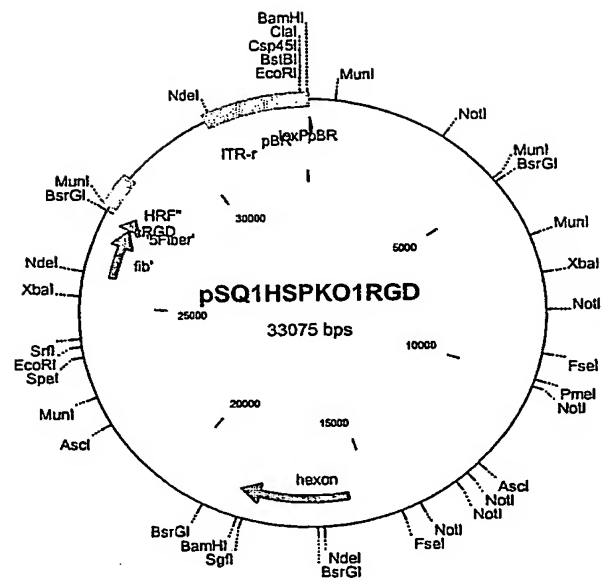


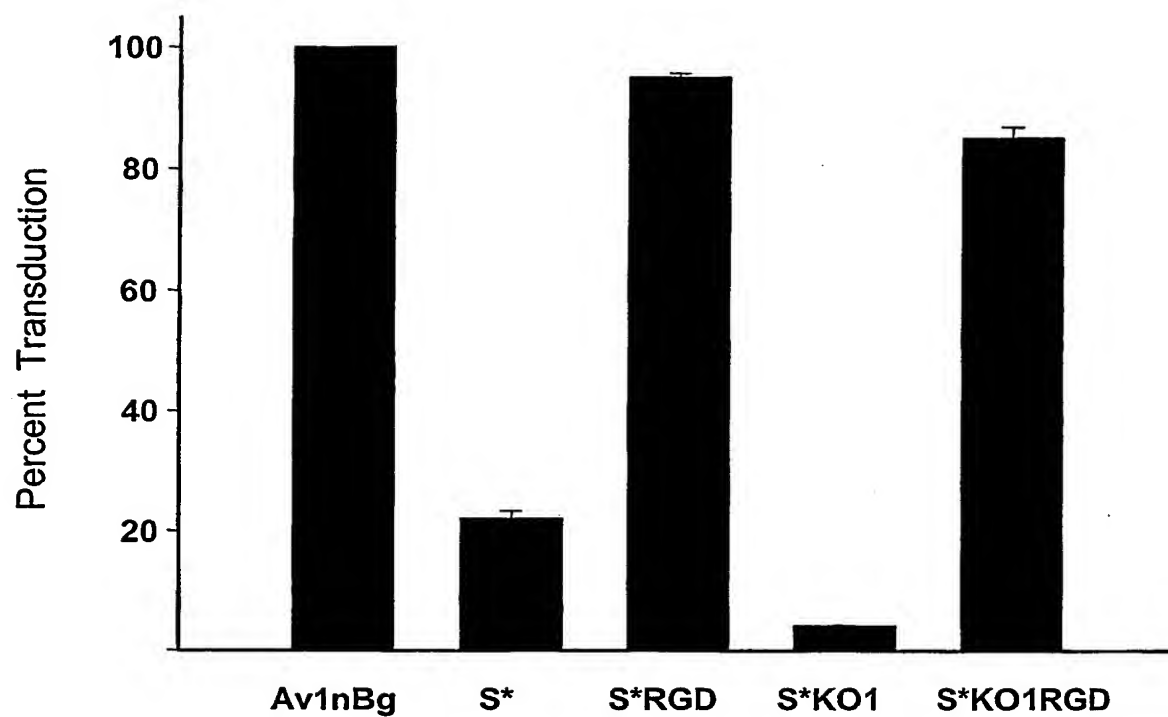
Figure 16

Figure 17A

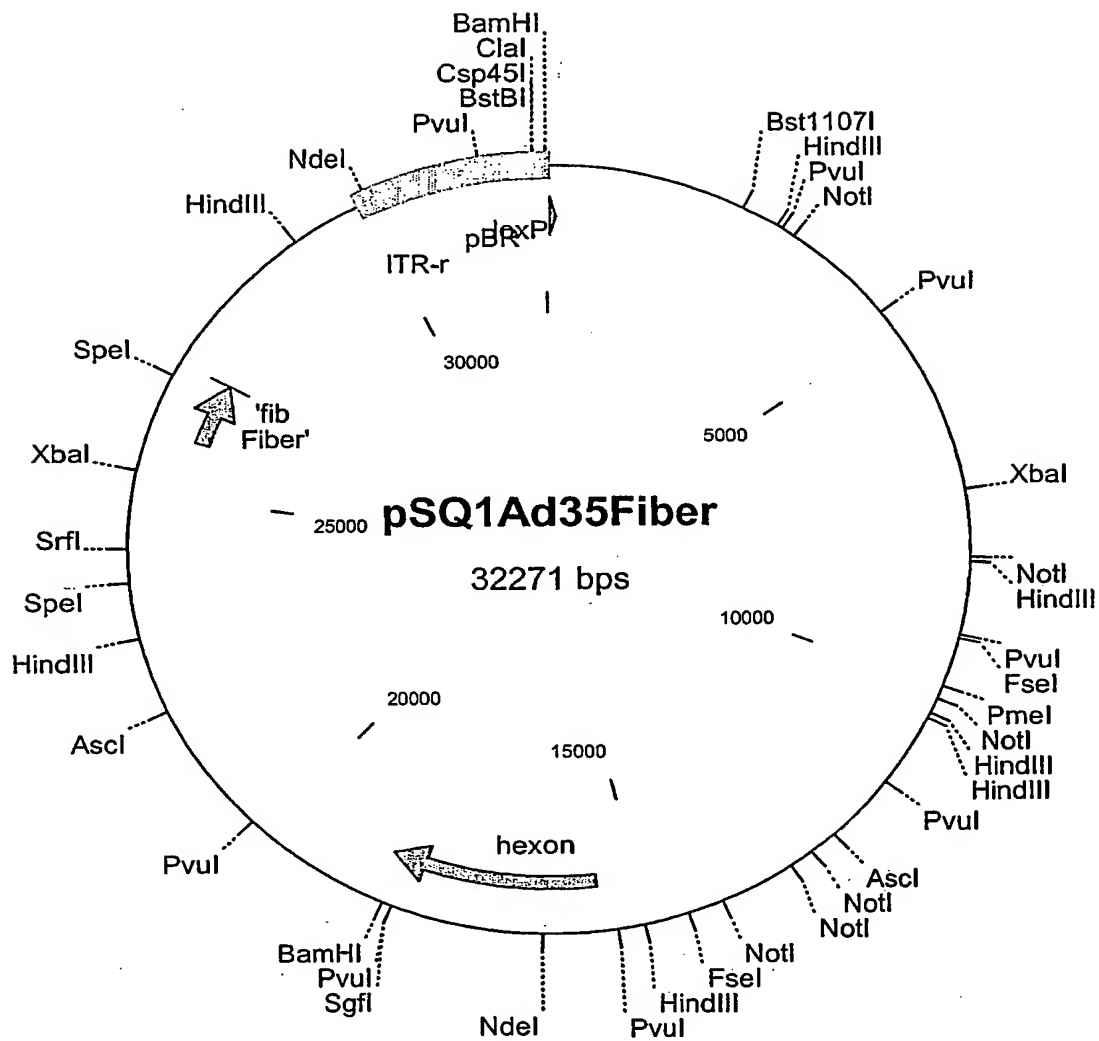


Figure 17B

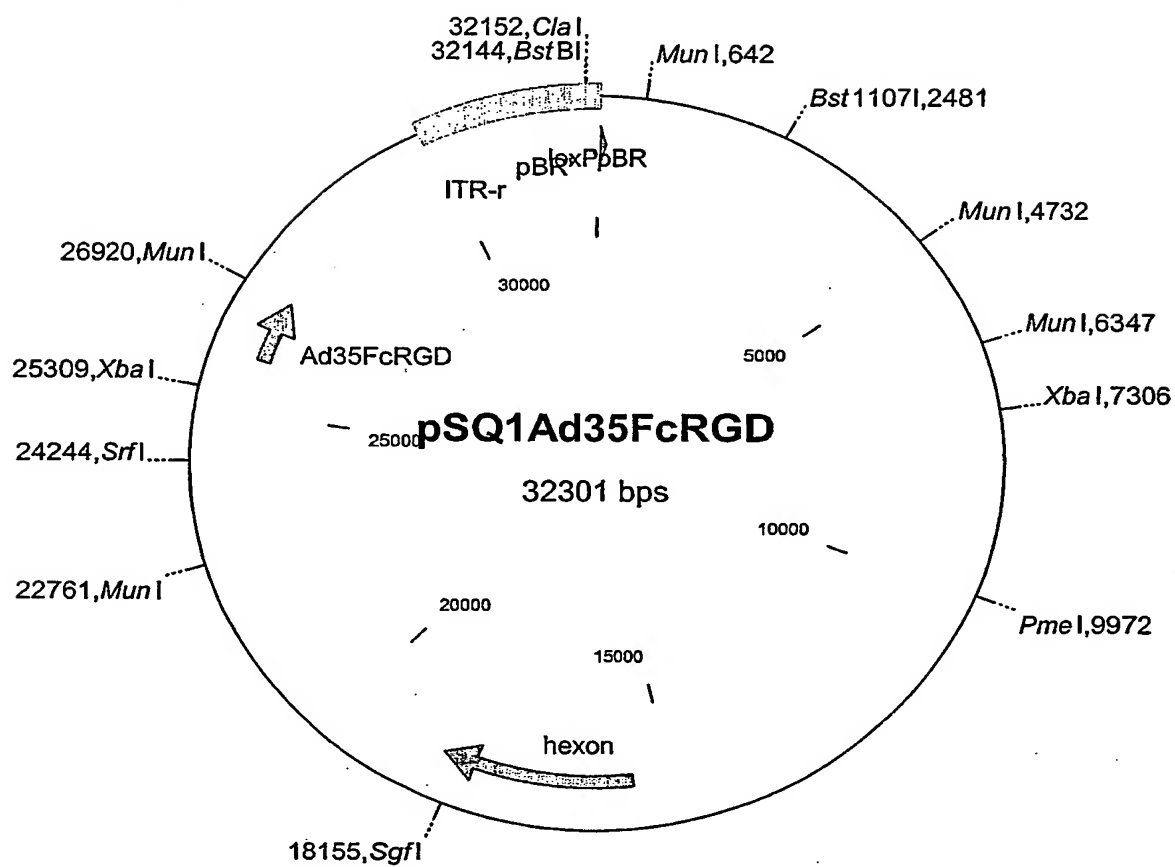


Figure 18A

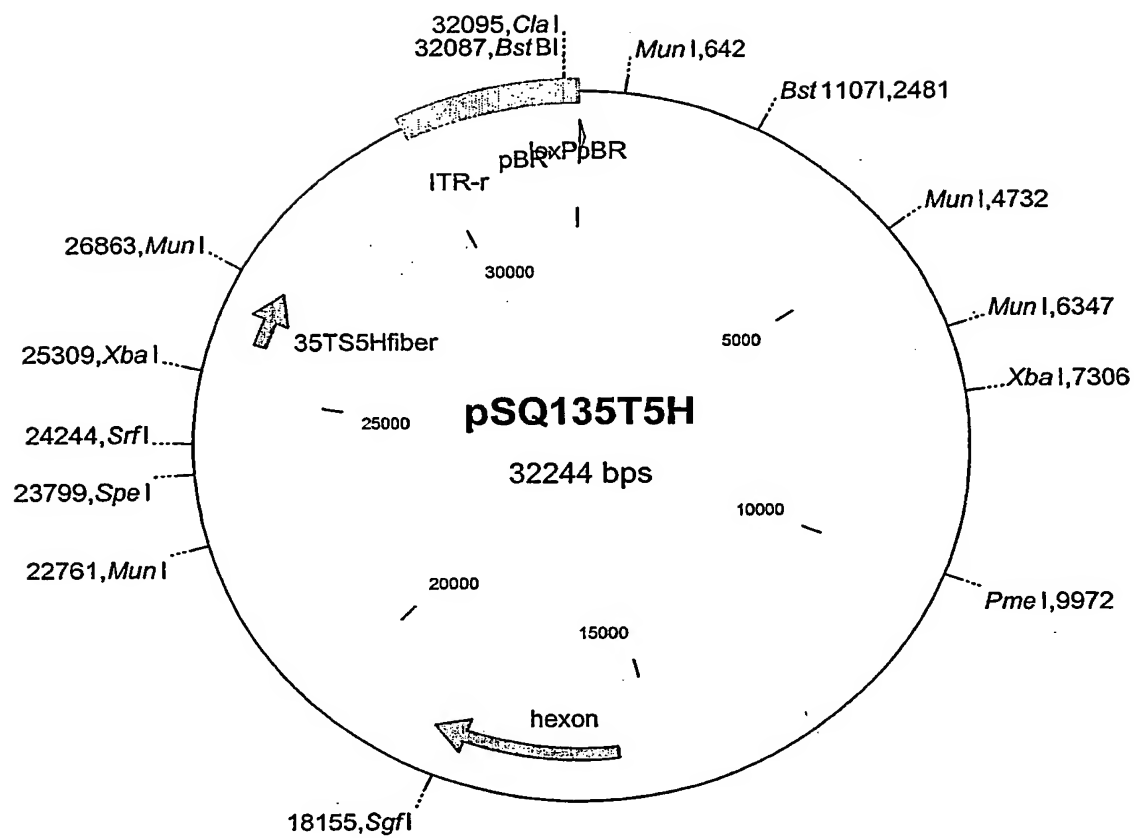


Figure 18B

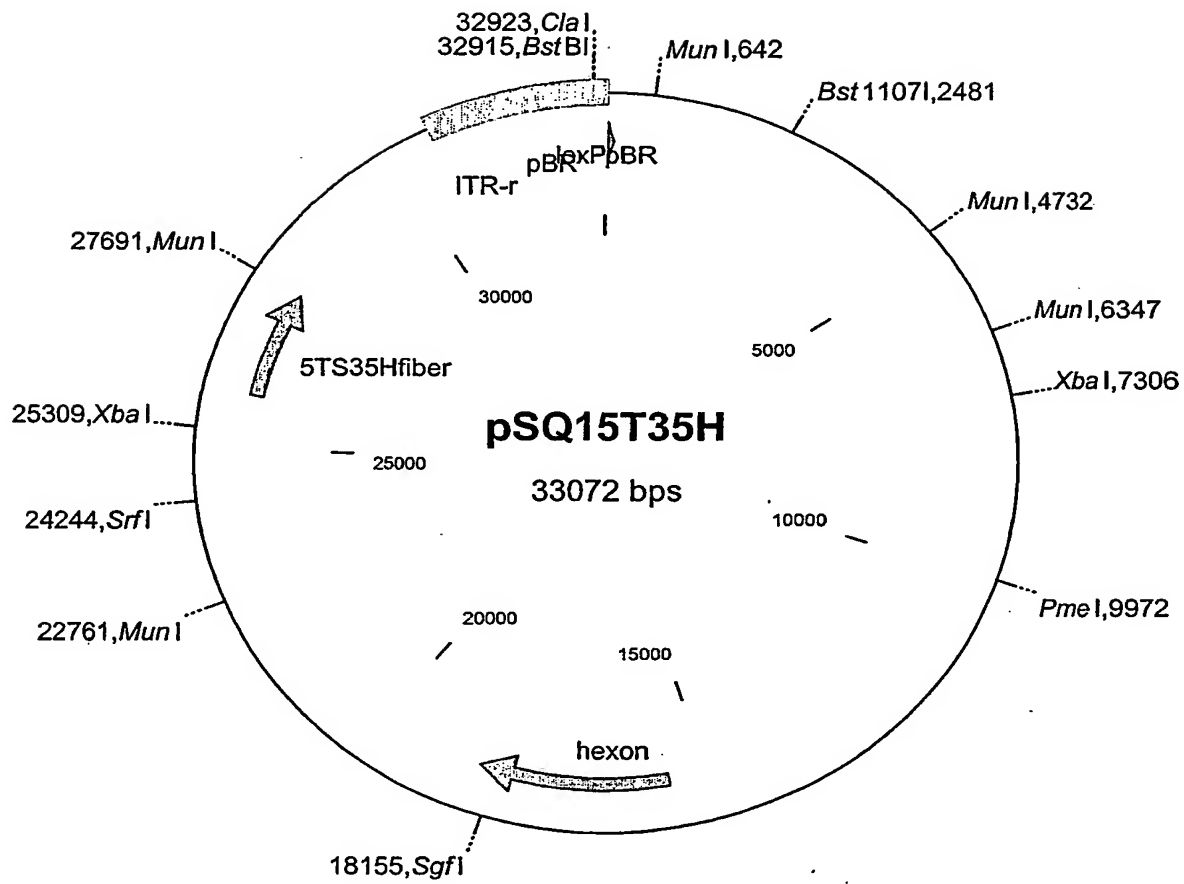


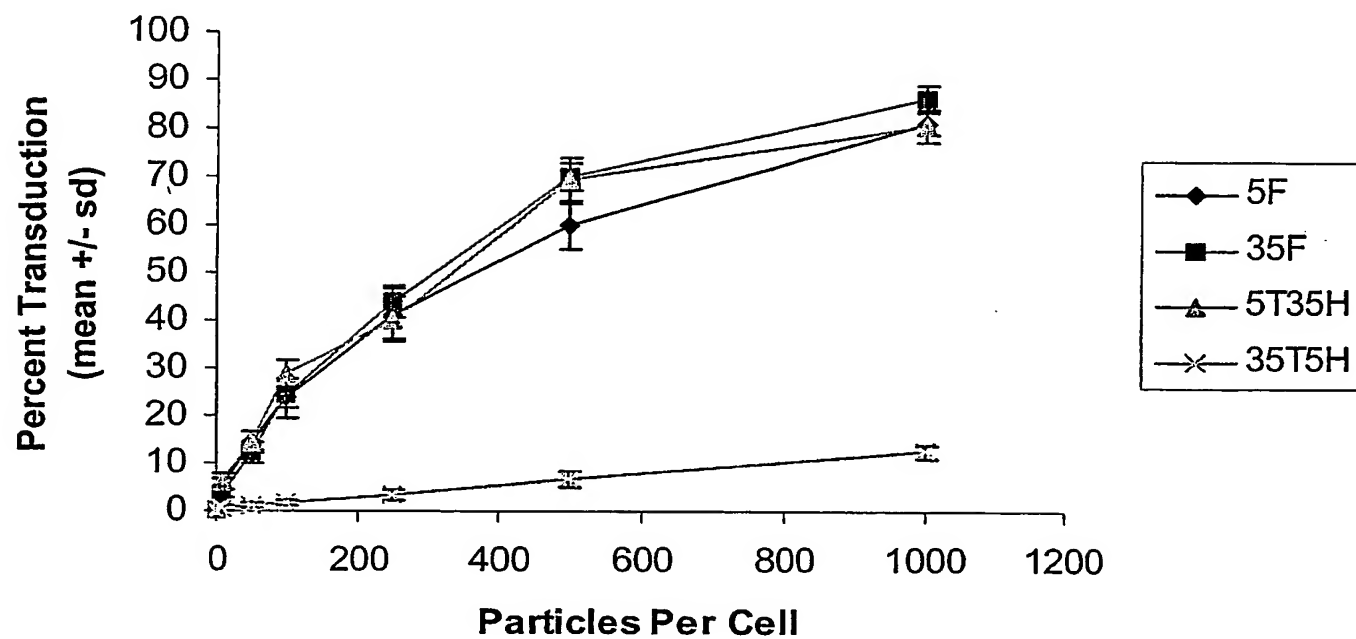
Figure 19

Figure 20

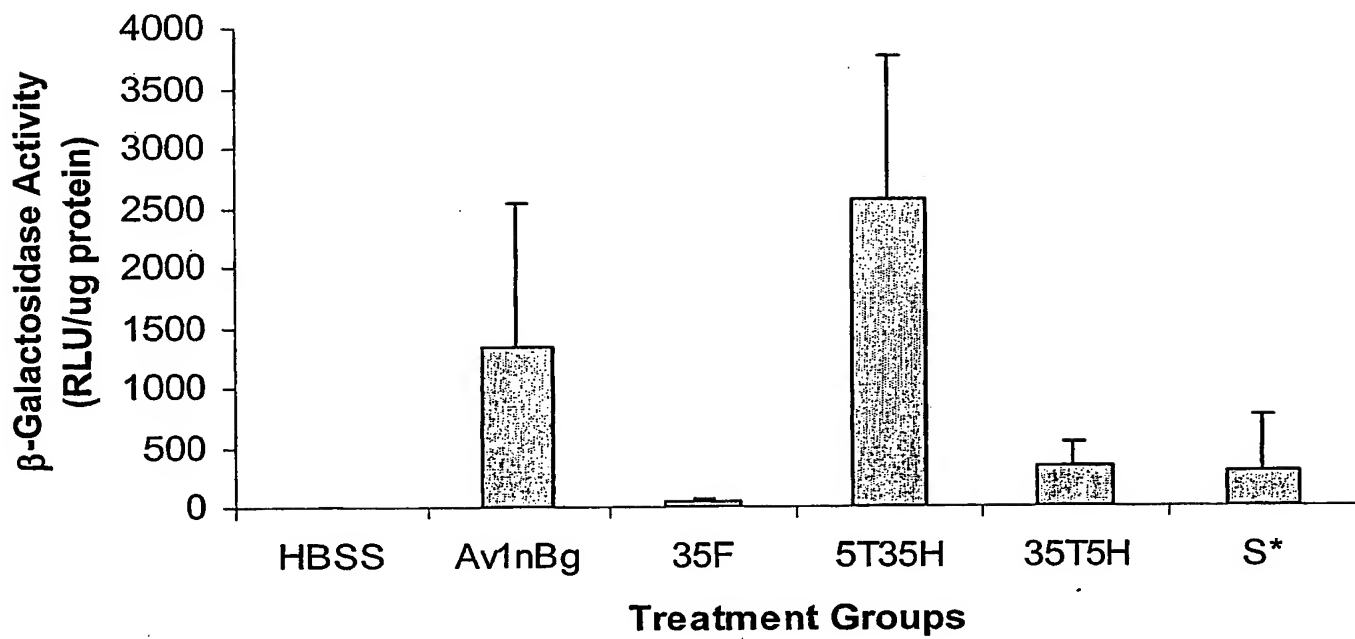


Figure 21A

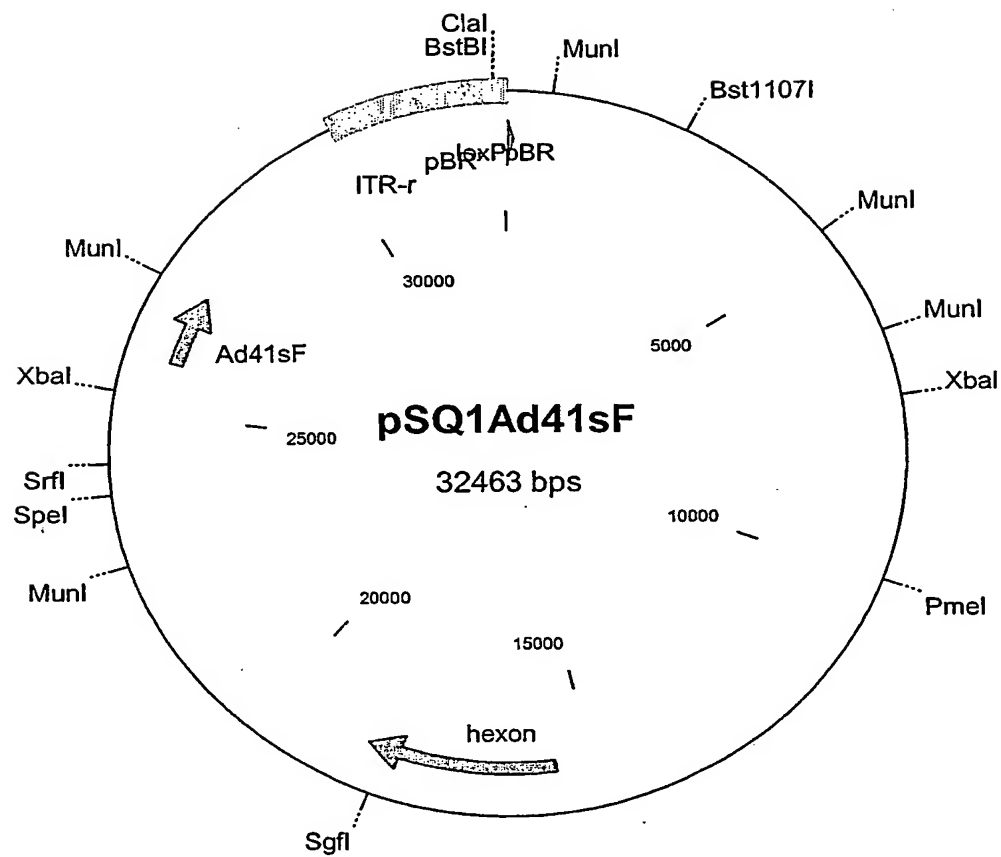
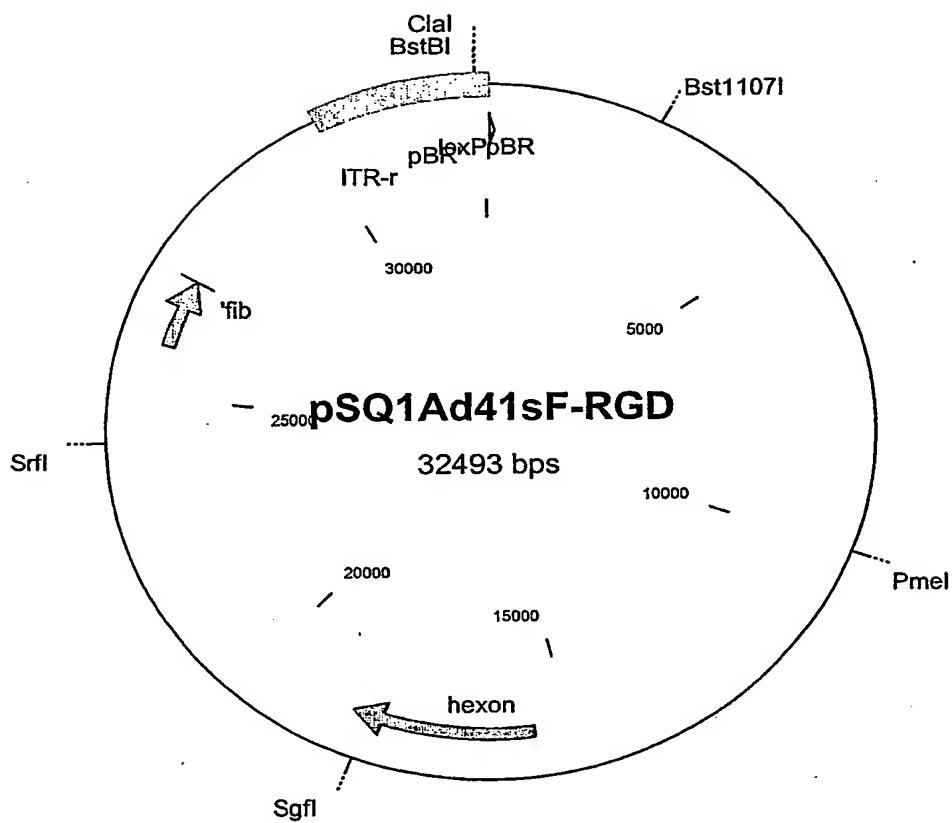


Figure 21B



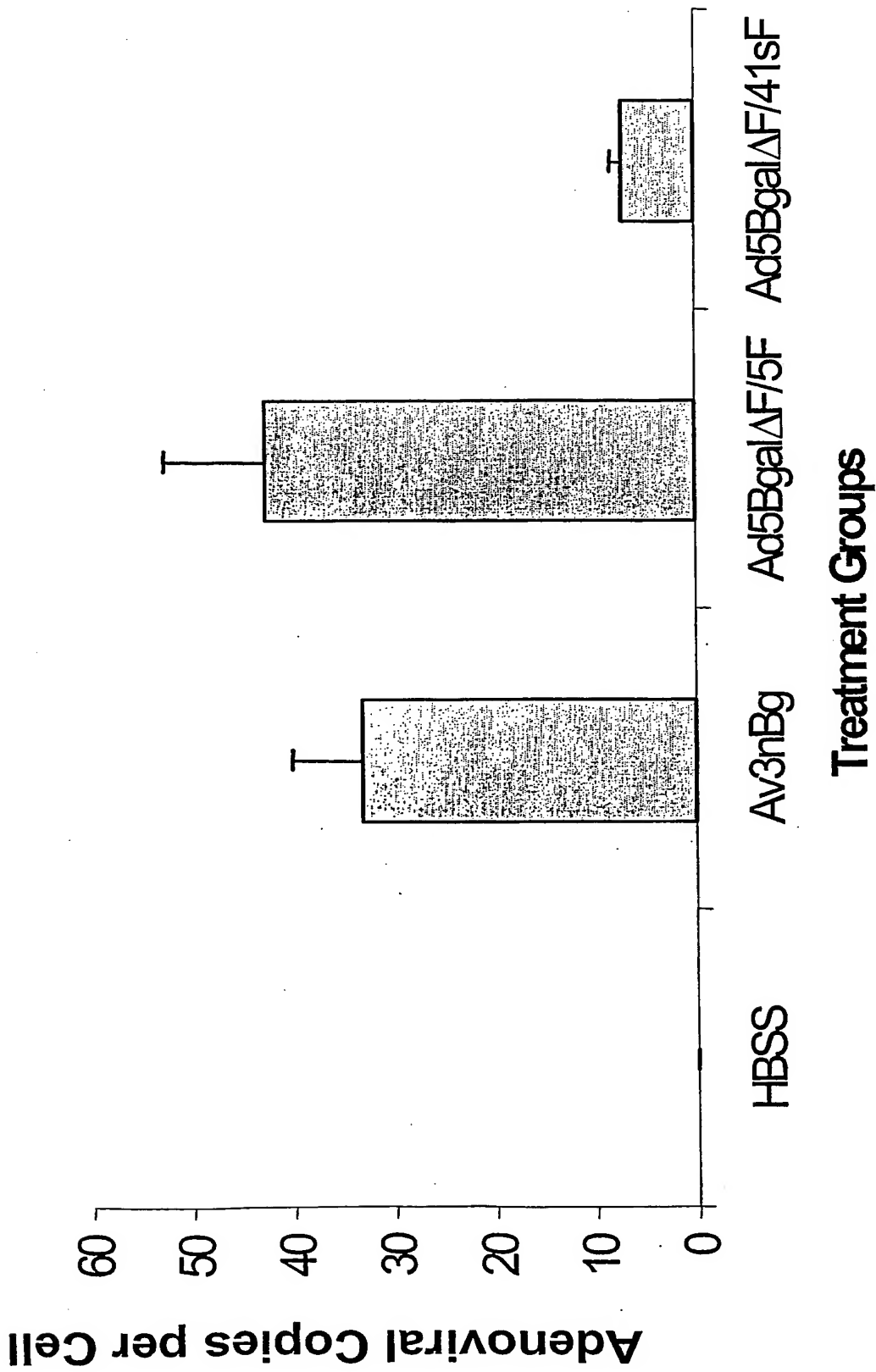


Figure 22

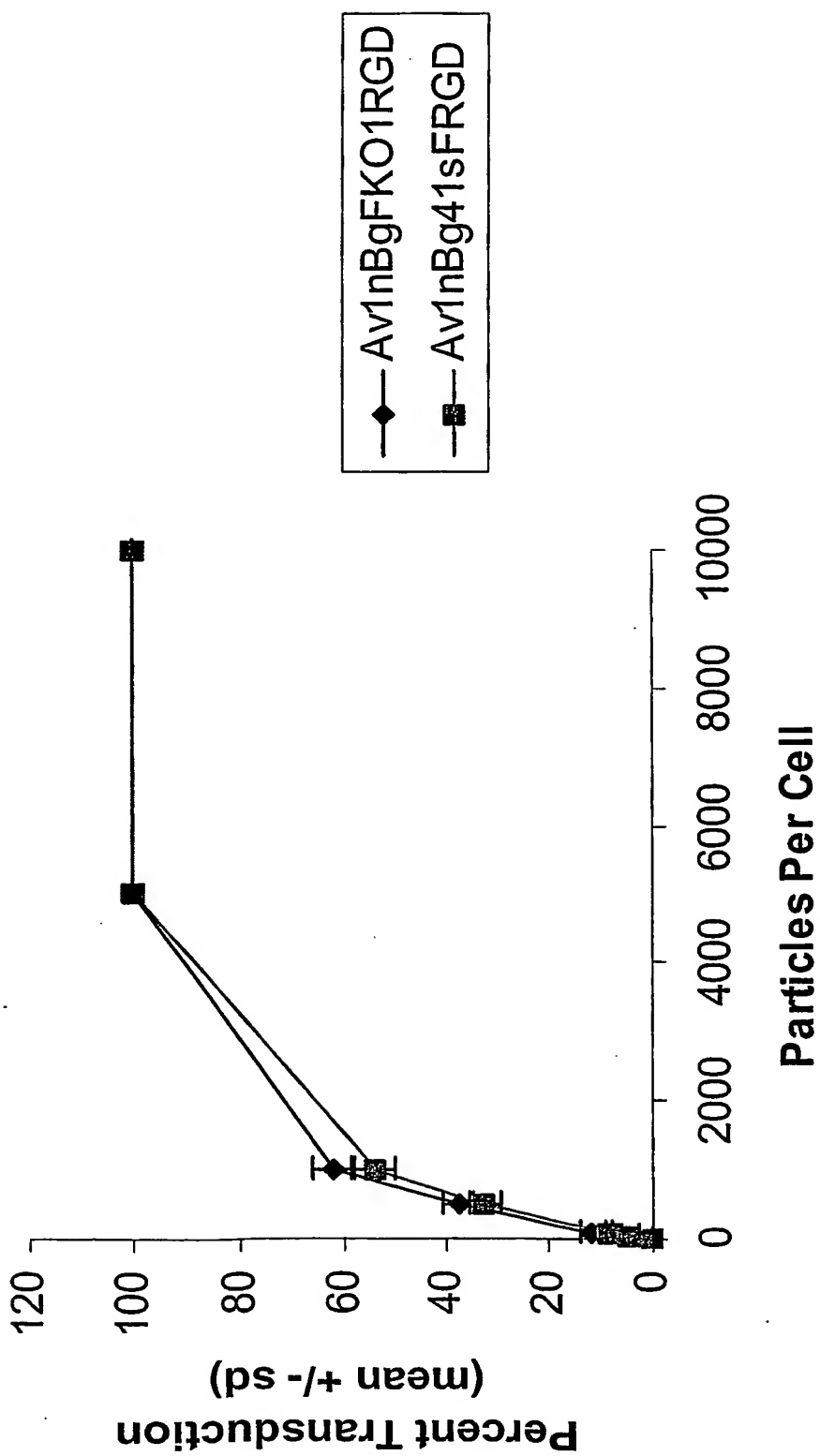
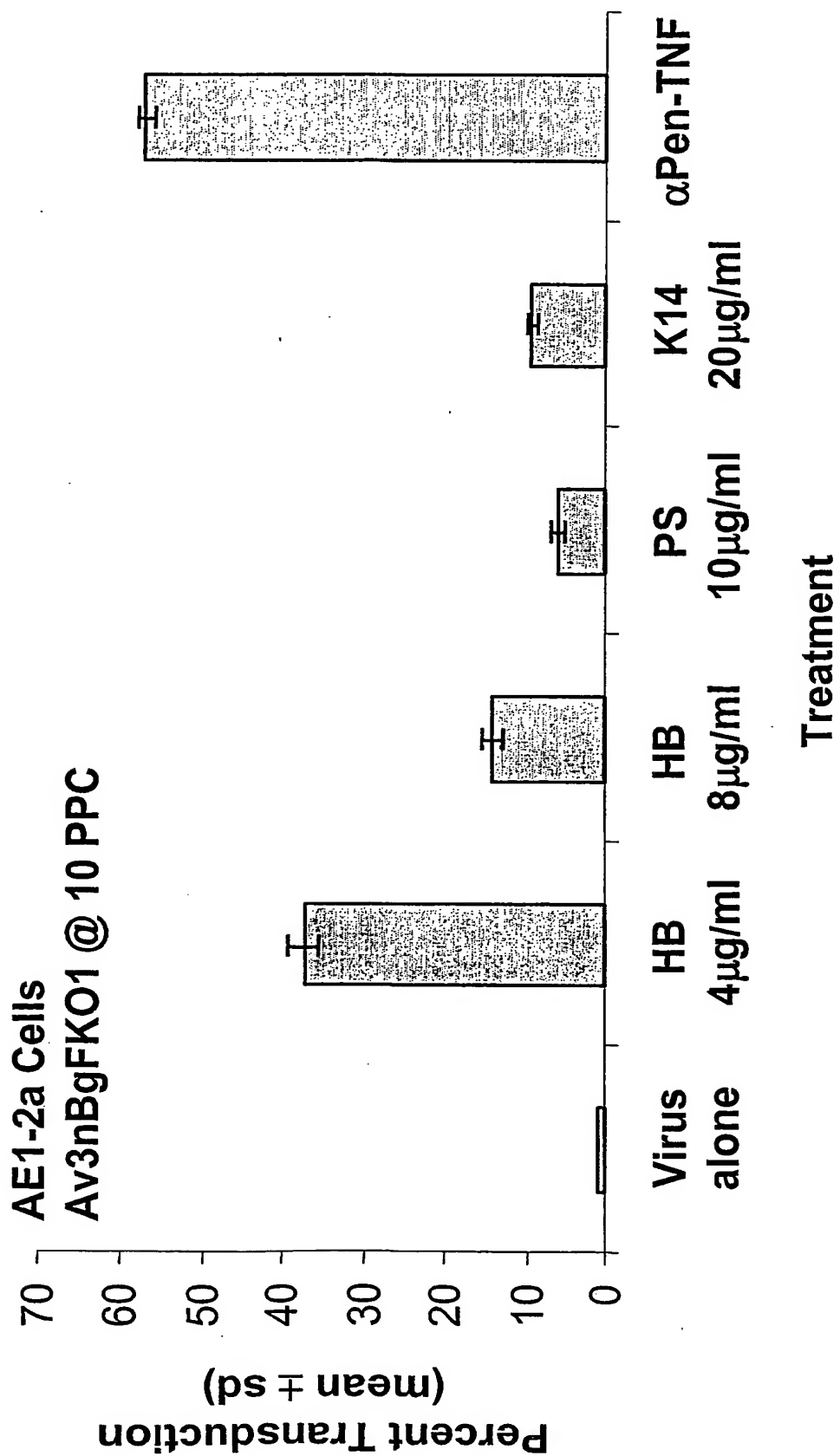


Figure 23

Figure 24



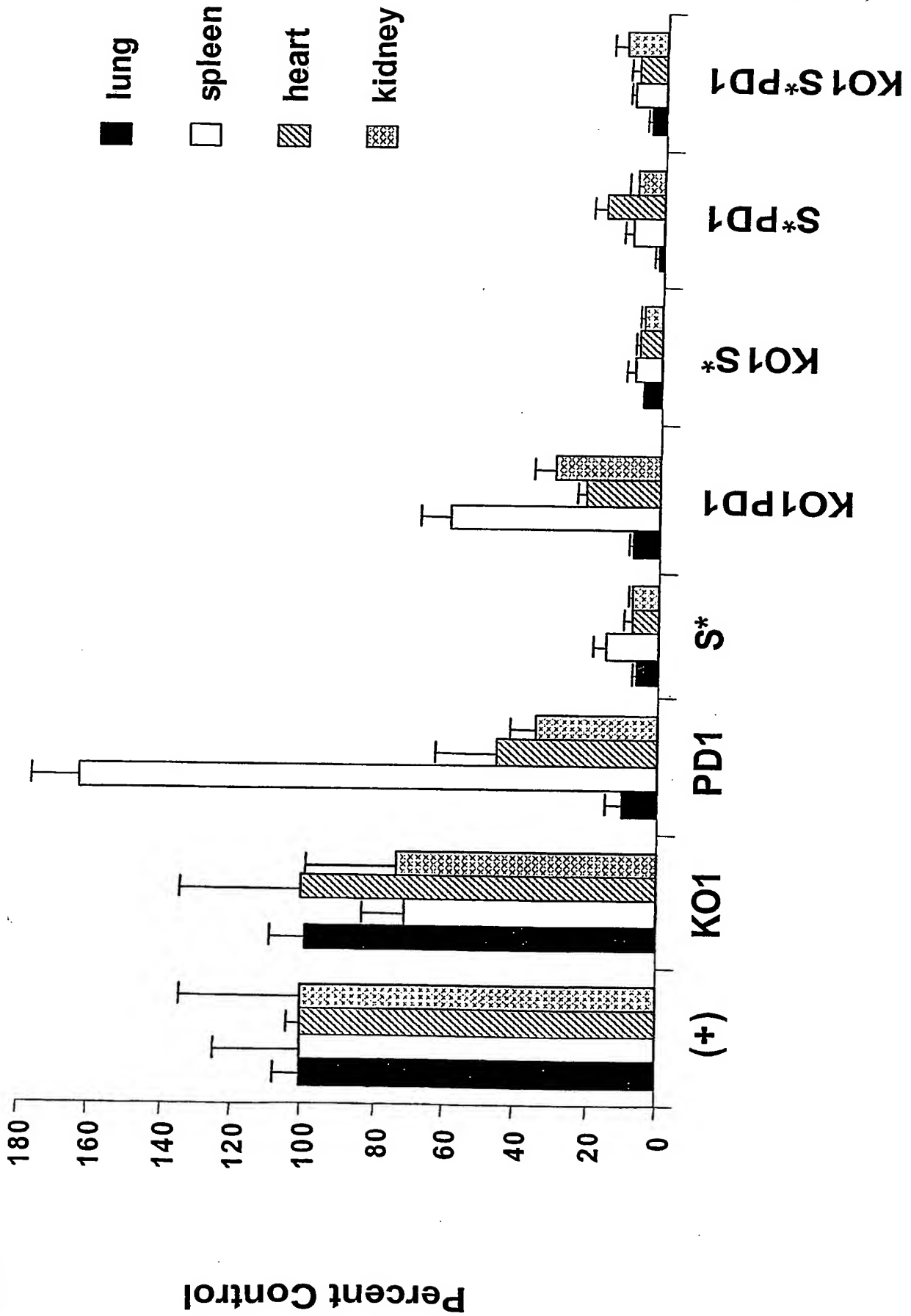


Figure 25

Figure 26

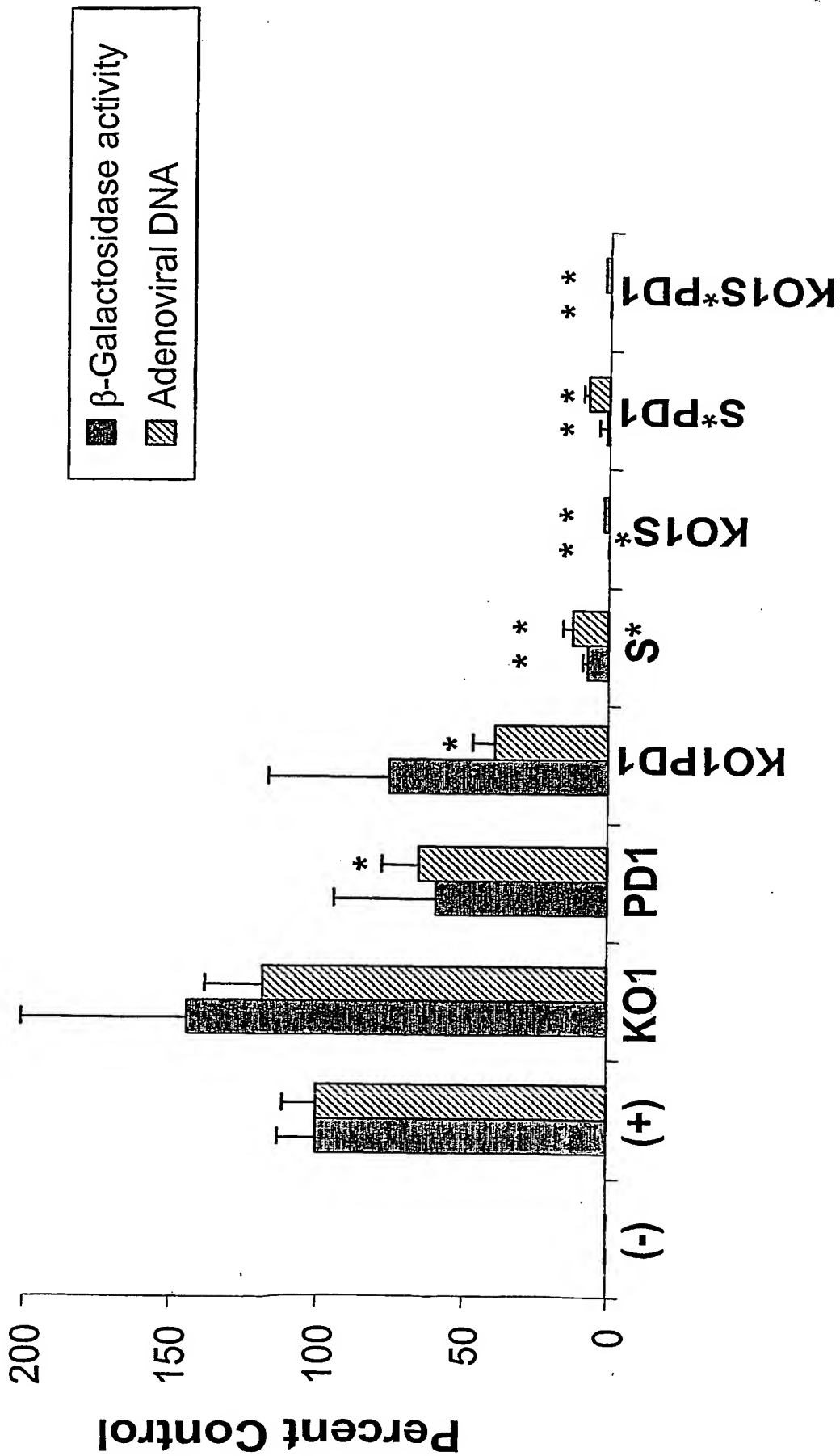
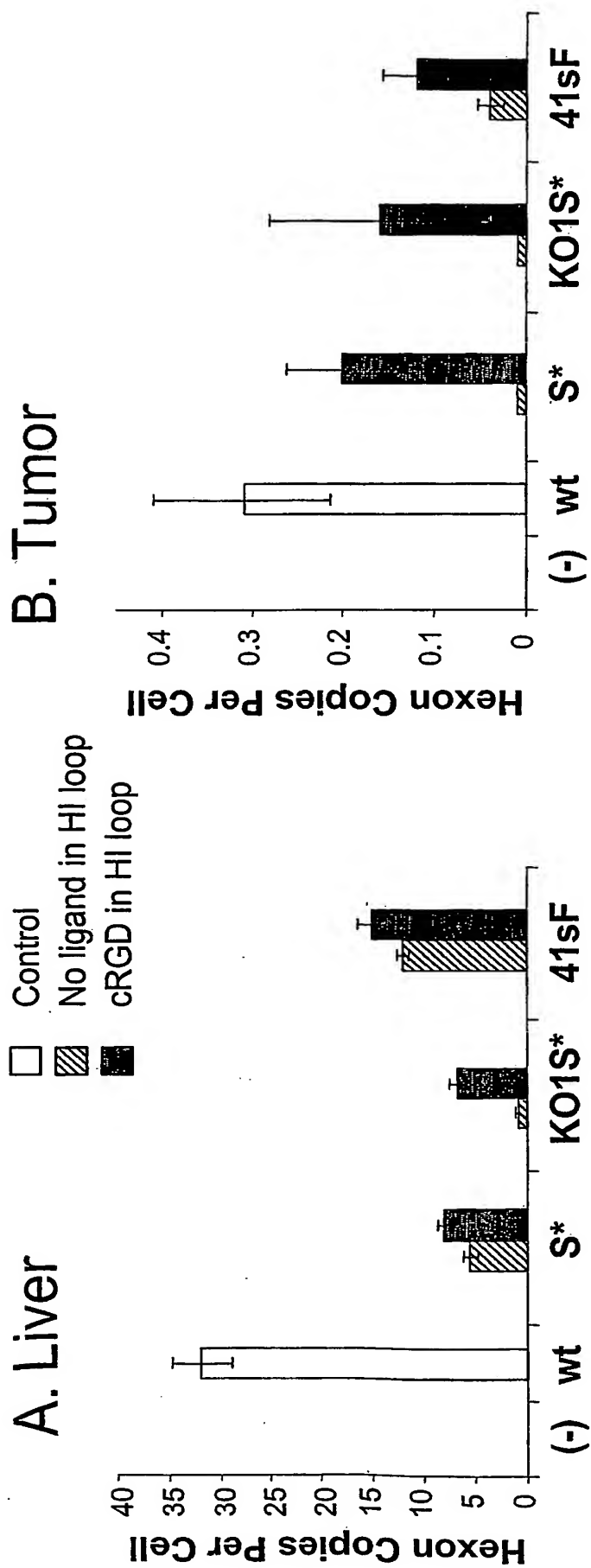


Figure 27



-1-

SEQUENCE LISTING

<110> The Scripps Research Institute
 Novartis AG
 Kaleko, Michael
 Nemerow, Glen R.
 Smith, Theodore
 Stevenson, Susan C.

<120> Fiber Shaft Modifications for Efficient Targeting

<130> 22908-1236PC

<140> Not yet assigned

<141> Herewith

<150> 60/350,388

<151> 2002-01-24

<150> 60/391,967

<151> 2002-06-26

<160> 72

<170> PatentIn version 3.0

<210> 1

<211> 4

<212> PRT

<213> adenovirus serotype 5

<400> 1
 Lys Lys Thr Lys
 1

<210> 2

<211> 1746

<212> DNA

<213> adenovirus serotype 5

<400> 2

atgaagcgcg	caagaccgtc	tgaagatacc	ttcaaccccg	tgtatccata	tgacacggaa	60
accggtcctc	caactgtgcc	ttttcttact	cctccctttg	tatcccccaa	tggttttcaa	120
gagagtcctc	ctgggggtact	ctctttgcgc	ctatccgaac	ctctagttac	ctccaatggc	180
atgcttgccg	tcaaaatggg	caacggcctc	tctctggacg	aggccggcaa	ccttacctcc	240
caaaatgtaa	ccactgtgag	cccacctctc	aaaaaaaacca	agtcaaacad	aaacctggaa	300
atatctgcac	ccctcacagt	tacctcagaa	gccctaactg	tgggtgccgc	cgcacctcta	360
atggctgcgg	gcaacacact	caccatgcaa	tcacaggccc	cgctaaccgt	gcacgactcc	420
aaacttagca	ttgccaccca	aggaccctc	acagtgtcag	aaggaaagct	agccctgcaa	480
acatcaggcc	ccctcaccac	caccgatagc	agtaccctta	ctatcactgc	ctcaccctcc	540
ctaactactg	ccactggtag	cttgggcatt	gacttgaaag	agcccattha	tacacaaaat	600
ggaaaactag	gactaaagta	cggggctcct	ttgcatgtaa	cagacgacct	aaacactttg	660
accgtagcaa	ctgggtccagg	tgtgactatt	aataatactt	ccttgcaaac	taaagttact	720
ggagccttgg	gtttttgattc	acaaggcaat	atgcaactta	atgtagcagg	aggactaagg	780
attgattctc	aaaacagacg	ccttatactt	gatgttagtt	atccgtttga	tgctcaaaaac	840
caactaaatc	taagactagg	acaggggcct	ctttttataa	actcagccca	caacttggat	900
attaactaca	acaaaggcct	ttacttgttt	acagcttcaa	acaattccaa	aaagcttgag	960
gttaacctaa	gcactgccaa	gggggtgatg	tttgacgcta	cagccatagc	cattaatgca	1020
ggagatgggc	ttgaatttgg	ttcacctaata	gcaccaaaca	caaatccctc	caaaacaaaa	1080
attggccatg	gcctagaatt	tgattcaaac	aaggctatgg	ttcctaaact	aggaactggc	1140
cttagttttg	acagcacagg	tgccattaca	gtaggaaaca	aaaataatga	taagctaact	1200
ttgtggacca	caccagctcc	agaggctaac	tgtagactaa	atgcagagaa	agatgctaaa	1260

-2-

ctcacttttg	tcttaacaaa	atgtggcagt	caaatacttg	ctacagtttc	agttttggct	1320
gttaaaggca	gtttggctcc	aatatctgga	acagttcaaa	gtgctcatct	tattataaga	1380
tttgacgaaa	atggagtgtc	actaaacaat	tccttctcgg	accagaata	ttggaacttt	1440
agaaatggag	atcttactga	aggcacagcc	tatacaaacg	ctgttggatt	tatgcctaac	1500
ctatcagctt	atccaaaatc	tcacggtaaa	actgcaaaa	gtaacattgt	cagtcaagtt	1560
tacttaaacy	gagacaaaac	taaacctgta	acactaacca	ttacactaaa	cggtacacag	1620
gaaacaggag	acacaactcc	aagtgcatac	tctatgtcat	tttcatggga	ctggtctggc	1680
cacaactaca	ttaatgaaat	atttgccaca	tcctcttaca	ctttttcata	cattgcccac	1740
gaataa						1746

<210> 3
 <211> 1746
 <212> DNA
 <213> adenovirus serotype 5

<400> 3						
atgaagcgcg	caagaccgtc	tgaagatacc	ttcaaccccg	tgtatccata	tgacacggaa	60
accggtcctc	caactgtgcc	ttttcttact	cctccctttg	tatcccccaa	tgggtttcaa	120
gagagtcccc	ctgggggtact	ctctttgctc	ctatccgaac	ctctagttac	ctccaatggc	180
atgcttgccg	tcaaaatggg	caacggcctc	tctctggacg	aggccggcaa	ccttacctcc	240
caaaatgtaa	ccactgtgag	cccacctctc	aaaaaaacca	agtcaaacat	aaacctggaa	300
atatctgcac	ccctcacagt	tacctcagaa	gccctaactg	tggctgccgc	cgcacctcta	360
atggctgcgg	gcaacacact	caccatgcaa	tcacaggccc	cgctaaccgt	gcacgactcc	420
aaacttagca	ttgccaccca	aggacccctc	acagtgtcag	aaggaaagct	agccctgcaa	480
acatcaggcc	ccctcaccac	caccgatagc	agtaccctta	ctatcactgc	ctcacccctc	540
ctaactactg	ccactggtag	cttggggcatt	gtgcatgtaa	agcccattta	tacacaaaat	600
ggaaaactag	gactaaagta	cggggctcct	ttgcatgtaa	cagacgacct	aaacactttg	660
accgtagcaa	ctgggtccagg	tgtgactatt	aataatactt	ccttgcaaac	taaagttact	720
ggagccttgg	gtttttgattc	acaaggcaat	atgcaactta	atgtagcagg	aggactaagg	780
attgattctc	aaaacagacg	ccttataactt	gatgttagtt	atccgtttga	tgctcaaaaac	840
caactaaatc	taagactagg	acagggccct	ctttttataa	actcagccca	caacttggat	900
attaactaca	acaaaggcct	ttacttgttt	acagcttcaa	acaattccaa	aaagcttgag	960
gttaacctaa	gcactgccaa	gggggttgatg	tttgacgcta	cagccatagc	cattaatgca	1020
ggagatgggc	ttgaatttgg	ttcacctaatt	gcaccaaaca	caaattccct	caaaacaaaa	1080
attggccatg	gcttagaatt	tgattcaaac	aaggctatgg	ttcctaaact	aggaactggc	1140
cttagttttg	acagcagagg	tgccattaca	gtaggaaaca	aaaataatga	taagctaact	1200
ttgtggacca	caccagctcc	atctcctaac	tgttcactaa	atggaggcgg	agatgctaaa	1260
ctcacttttg	tcttaacaaa	atgtggcagt	caaatacttg	ctacagtttc	agttttggct	1320
gttaaaggca	gtttggctcc	aatatctgga	acagttcaaa	gtgctcatct	tattataaga	1380
tttgacgaaa	atggagtgtc	actaaacaat	tccttctcgg	accagaata	ttggaacttt	1440
agaaatggag	atcttactga	aggcacagcc	tatacaaacg	ctgttggatt	tatgcctaac	1500
ctatcagctt	atccaaaatc	tcacggtaaa	actgcaaaa	gtaacattgt	cagtcaagtt	1560
tacttaaacy	gagacaaaac	taaacctgta	acactaacca	ttacactaaa	cggtacacag	1620
gaaacaggag	acacaactcc	aagtgcatac	tctatgtcat	tttcatggga	ctggtctggc	1680
cacaactaca	ttaatgaaat	atttgccaca	tcctcttaca	ctttttcata	cattgcccac	1740
gaataa						1746

<210> 4
 <211> 1737
 <212> DNA
 <213> adenovirus serotype 5

<400> 4						
atgcggcgcg	cggcgatgta	tgagggaaggt	cctcctccct	cctacgagag	tgtgggtgagc	60
gcggcgccag	tggcgggcggc	gctgggttct	cccttcgatg	ctcccttgga	cccgcggttt	120
gtgcctccgc	ggtacctgcg	gcctaccggg	gggagaaaca	gcatccgtta	ctctgagttg	180
gcacccctat	tcgacaccac	ccgtgtgtac	ctgggtggaca	acaagtcaac	ggatgtggca	240
tccttgaact	accagaacga	ccacagcaac	tttctgacca	cggctcattca	aaacaatgac	300
tacagcccg	gggaggcaag	cacacagacc	atcaatcttg	acgaccggtc	gcactggggc	360
ggcgacctga	aaaccatcct	gcataccaac	atgccaaatg	tgaacgagtt	catgtttacc	420
aataagttta	aggcgcggggt	gatggtgtcg	cgcttgcccta	ctaaggacaa	tcaggtggag	480

2

<210>	5	
<211>	20	
<212>	DNA	
<213>	adenovirus serotype 5	
<400>	5	
	gaacaggagg tgagcttaga	20
<210>	6	
<211>	43	
<212>	DNA	
<213>	adenovirus serotype 5	
<400>	6	
	tccgcctcca tttagtgaac agttaggaga tggagctggt gtg	43
<210>	7	
<211>	44	
<212>	DNA	
<213>	adenovirus serotype 5	
<400>	7	
	tcactaaatg gaggcggaga tgctaaactc actttggtct taac	44
<210>	8	
<211>	20	
<212>	DNA	
<213>	adenovirus serotype 5	
<400>	8	
	gtggcagggt gaatactagg	20
<210>	9	
<211>	8	
<212>	PRT	
<213>	adenovirus serotype 5	

<400> 9
His Ala Ile Arg Gly Asp Thr Phe

-4-

1 5

<210> 10
 <211> 15
 <212> PRT
 <213> Artificial Sequence
 <220>
 <223> modified sequence for penton protein

<400> 10
 Ser Arg Gly Tyr Pro Tyr Asp Val Pro Asp Tyr Ala Gly Thr Ser
 1 5 10 15

<210> 11
 <211> 57
 <212> DNA
 <213> Artificial Sequence
 <220>
 <223> oligonucleotide for mutation generation

<400> 11
 cgcggaagag aactccaacg cggcagccgc ggcaatgcag ccggtggagg acatgaa 57

<210> 12
 <211> 59
 <212> DNA
 <213> Artificial Sequence
 <220>
 <223> oligonucleotide for mutation generation

<400> 12
 tatcggtcat gtctccacc ggctgcattg ccgcggctgc cgcgttggag ttctcttcc 59

<210> 13
 <211> 75
 <212> DNA
 <213> Artificial Sequence
 <220>
 <223> oligonucleotide for mutation generation

<400> 13
 cgatagccgc ggctaccct acgacgtgcc cgactacgcg ggcaccagcg ccacacgggc 60
 tgaggagaag cgcgc 75

<210> 14
 <211> 73
 <212> DNA
 <213> Artificial Sequence
 <220>
 <223> oligonucleotide for mutation generation

<400> 14
 tcagcgcgct tctcctcagc ccgtgtggcg ctggtgcccg cgtagtcggg cacgtcgtag 60
 gggtagccgc ggc 73

<210> 15
 <211> 40
 <212> DNA
 <213> Artificial Sequence
 <220>
 <223> oligonucleotide for mutation generation

-5-

<400> 15
 ggctccggct ccgagaggtg ggctcacagt ggttacattt 40
 <210> 16
 <211> 32
 <212> DNA
 <213> Artificial Sequence
 <220>
 <223> oligonucleotide for mutation generation

<400> 16
 ggagccggag cctcaaacaat aaacctggaa at 32
 <210> 17
 <211> 27
 <212> DNA
 <213> Artificial Sequence
 <220>
 <223> amplification primer

<400> 17
 ctctagaaat ggacggaatt attacag 27
 <210> 18
 <211> 32
 <212> DNA
 <213> Artificial Sequence
 <220>
 <223> amplification primer

<400> 18
 tcttggtcat ctgcaacaac atgaagatag tg 32
 <210> 19
 <211> 32
 <212> DNA
 <213> Artificial Sequence
 <220>
 <223> amplification primer

<400> 19
 gttgttgtag atgaccaaga gagtccggct ca 32
 <210> 20
 <211> 73
 <212> DNA
 <213> Artificial Sequence
 <220>
 <223> amplification primer

<400> 20
 agcaattgaa aaataaacac gttgaaacat aacacaaacg attcttttagt tgtcgtcttc 60
 tgtaatgtaa gaa 73
 <210> 21
 <211> 24
 <212> DNA
 <213> Artificial Sequence
 <220>
 <223> amplification primer

<400> 21

-6-

agcaattgaa aaataaacac gttg 24

<210> 22
<211> 20
<212> DNA
<213> Artificial Sequence
<220>
<223> amplification primer

<400> 22
gaacaggagg tgagcttaga 20

<210> 23
<211> 42
<212> DNA
<213> Artificial Sequence
<220>
<223> amplification primer

<400> 23
gttaggtgga ggggtttattc cgggtccacaa agttagctta tc 42

<210> 24
<211> 42
<212> DNA
<213> Artificial Sequence
<220>
<223> amplification primer

<400> 24
gataagctaa ctttgtggac cggaataaac cctccaccta ac 42

<210> 25
<211> 20
<212> DNA
<213> Artificial Sequence
<220>
<223> amplification primer

<400> 25
gtggcagggt gaatactagg 20

<210> 26
<211> 41
<212> DNA
<213> Artificial Sequence
<220>
<223> amplification primer

<400> 26
gttaggagat ggagctggtg tagtccataa ggtggttaata c 41

<210> 27
<211> 41
<212> DNA
<213> Artificial Sequence
<220>
<223> amplification primer

<400> 27
gtattaacac cttatggact acaccagctc catctcctaa c 41

-7-

<210> 28
 <211> 54
 <212> DNA
 <213> Artificial Sequence
 <220>
 <223> amplification primer

<400> 28
 tgcgcaaaaa caatcaccac gacaatcaca atgtacattg gaagaaatca tacg 54

<210> 29
 <211> 54
 <212> DNA
 <213> Artificial Sequence
 <220>
 <223> amplification primer

<400> 29
 acattgtgat tgcgtggtg attgtttttg cgcatatgcc atacaatttg aatg 54

<210> 30
 <211> 10
 <212> PRT
 <213> Artificial Sequence
 <220>
 <223> RGD targeting peptide

<400> 30
 His Cys Asp Cys Arg Gly Asp Cys Phe Cys
 1 5 10

<210> 31
 <211> 32
 <212> DNA
 <213> Artificial Sequence
 <220>
 <223> amplification primer

<400> 31
 ttcttttcat ctgcaacaac atgaagatag tg 32

<210> 32
 <211> 32
 <212> DNA
 <213> Artificial Sequence
 <220>
 <223> amplification primer

<400> 32
 gttgttgcag atgaaaagaa ccagaattga ag 32

<210> 33
 <211> 73
 <212> DNA
 <213> Artificial Sequence
 <220>
 <223> amplification primer

<400> 33
 tgcaattgaa aaataaacac gttgaaacat aacacaaacg attctttatt cttcagttat 60
 gtagcaaaat aca 73

-8-

<210> 34
<211> 56
<212> DNA
<213> Artificial Sequence
<220>
<223> amplification primer

<400> 34
agtacaaaaa caatcaccac gacaatcaca gtttatctcg ttgtagacga cactga 56

<210> 35
<211> 51
<212> DNA
<213> Artificial Sequence
<220>
<223> amplification primer

<400> 35
tgtgattgtc gtggtgattg tttttgtact agtgggtatg cttttacttt t 51

<210> 36
<211> 4
<212> PRT
<213> Adenovirus type 5

<400> 36
Thr Leu Trp Thr
1

<210> 37
<211> 7
<212> PRT
<213> SV40

<400> 37
Pro Lys Lys Lys Arg Lys Val
1 5

<210> 38
<211> 19
<212> DNA
<213> Artificial Sequence
<220>
<223> amplification primer

<400> 38
cttcgatgat gccgcagtg 19

<210> 39
<211> 19
<212> DNA
<213> Artificial Sequence
<220>
<223> amplification primer

<400> 39
gggctcaggt actccgagg 19

<210> 40
<211> 25
<212> DNA

-9-

<213> Artificial Sequence
 <220>
 <223> amplification primer

<400> 40
 ttacatgcac atctcgggcc aggac

25

<210> 41
 <211> 7607
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Plasmid GRE5-E1-SV40-Hygro

<400> 41
 tctagaagat ccgctgtaca ggatgttcta gctactttat tagatccgct gtacaggatg 60
 ttctagctac tttattagat ccgctgtaca ggatgttcta gctactttat tagatccgct 120
 gtacaggatg ttctagctac tttattagat ccgctgtacag gatgttctag ctactttatt 180
 agatcgatct cctggccggt cgggggtcaaa aaccagggtt ggctataaaa ggggggtgggg 240
 gcgcgttcgt cctcactctc ttccgcacgc ctgtctgcga gggccaggat cgatcctgag 300
 aacttcagggt tgagtttggg gacccttgat tgttctttct ttttcgctat tgtaaaattc 360
 atgttatatg gagggggcaa agttttcagg gtgttgttta gaatgggaag atgtcccttg 420
 tatcaccatg gaccctcatg ataattttgt ttctttcact ttctactctg ttgacaacca 480
 ttgtctcctc ttatttttct ttcattttct gtaacttttt cgttaaactt tagcttgcat 540
 ttgtaacgaa tttttaaatt cacttttgtt tatttgtcag attgtaagta ctttctctaa 600
 tcactttttt ttcaaggcaa tcagggtata ttatattgta cttcagcaca gtttttagaga 660
 acaattgtta taattaaatg ataaggtaga atattttctgc atataaattc tggctggcgt 720
 ggaaatatct ttattggtag aaacaactac atcctgggtc tcatcctgcc tttctcttta 780
 tggttacaat gatatacact gtttgagatg aggataaaat actctgagtc caaacccggc 840
 ccctctgcta accatgttca tgccttcttc tttttcctac agctcctggg caacgtgctg 900
 gttattgtgc tgtctcatca ttttggcaaa gaattagatc taagcttctg cagctcgagg 960
 actcggctga ctgaaaatga gacatattat ctgccacgga ggtgttatta ccgaagaaat 1020
 ggcgccaggt cttttggacc agctgatcga agaggtactg gctgataatc ttccacctcc 1080
 tagccatttt gaaccaccta cccttcacga actgtatgat ttagacgtga cggcccccga 1140
 agatcccaac gaggaggcgg tttcgcagat ttttcccgac tctgtaatgt tggcgggtgca 1200
 ggaagggtat gacttactca cttttccgcc ggcgcgccgt tctccggagc cgcctcacct 1260
 ttcccgagcag ccgagcagc cggagcagag agccttgggt ccggtttcta tgccaaacct 1320
 tgtaccggag gtgactgcac ttacctgcca cgaggtggc tttccacca gtgacgaca 1380
 ggatgaagag ggtgaggagt ttgtgttaga ttatgtggag caccgccggc acggttgca 1440
 gtcttgtcat tatcaccgga ggaatacggg ggaccagat attatgtgtt cgctttgcta 1500
 tatgaggacc tgtggcatgt ttgtctacag taagtgaata ttatgggcag tgggtgatag 1560
 agtgggtggg ttgggtgtgt aatttttttt ttaattttta cagttttgtg gtttaagaa 1620
 ttttgtattg tgattttttt aaaaggctct gtgtctgaac ctgagcctga gcccgagcca 1680
 gaaccggagc ctgcaagacc taccgcctgt cctaaaatgg cgcctgctat cctgagacgc 1740
 ccgacatcac ctgtgtctag agaatgcaat agtagtacgg atagctgtga ctccggtcct 1800
 tctaacacac ctctgagat acaccgggtg gtcccgctgt gccccattaa accagttgcc 1860
 gtgagagtgt gtgggcgtcg ccaggctgtg gaatgtatcg aggacttgct taacgagcct 1920
 gggcaacctt tggacttgag ctgtaaacgc ccaggccat aagggtgtaa cctgtgattg 1980
 cgtgtgtggg taacgccttt gtttgctgaa tgagttgatg taagttaaat aaagggtgag 2040
 ataattgtta acttgcatgg cgtgttaaat ggggcggggc ttaaagggtg tataatgcgc 2100
 cgtgggctaa tcttgggtac atctgacctc atggaggctt gggagtgtt ggaagatttt 2160
 tctgctgtgc gtaacttgct ggaacagagc tctaacagta cctcttggtt ttggaggttt 2220
 ctgtggggct catcccaggc aaagttagtc tgcagaatta aggaggatta caagtgggaa 2280
 tttgaagagc ttttgaaatc ctgtggtgag ctgtttgatt ctttgaatct gggtcaccag 2340
 gcgcttttcc aagagaagggt catcaagact ttggattttt ccacaccggg gcgcgctgcg 2400
 gctgctgttg cttttttgag gtaacttgct gataaattgga gcgaagaaac ccatctgagc 2460
 ggggggtacc tgctggattt tctggccatg catctgtgga gagcgggtgt gagacacaag 2520
 aatcgccctgc tactgttgct ttccgtccgc ccggcgataa taccgacgga ggagcagcag 2580
 cagcagcagg aggaagccag gcggcgccgc caggagcaga gcccatggaa cccgagagcc 2640
 ggcctggacc ctccgggaatg aatgttgtac aggtggctga actgtatcca gaactgagac 2700

-10-

gcattttgac	aattacagag	gatgggcagg	ggctaaagg	ggtaaagagg	gagcgggggg	2760
cttgtgaggc	tacagaggag	gctaggaatc	tagcttttag	cttaatgacc	agacaccgctc	2820
ctgagtgtat	tacttttcaa	cagatcaagg	ataattgcgc	taatgagctt	gatctgctgg	2880
cgcagaagta	ttccatagag	cagctgacca	cttactggct	gcagccagg	gatgattttg	2940
aggaggctat	taggggtatat	gcaaagggtg	cacttaggcc	agattgcaag	tacaagatca	3000
gcaaacttgt	aaatatcagg	aattgttgct	acatttctgg	gaacggggcc	gagggtggaga	3060
tagatacggg	ggatagggtg	gccttttagat	gtagcatgat	aaatatgtgg	ccgggggtgc	3120
ttggcatgga	cgggggtggt	attatgaatg	taagggtttac	tggccccaat	tttagcggta	3180
cggttttcct	ggccaatacc	aaccttatcc	tacacgggtg	aagcttctat	gggtttaaca	3240
atacctgtgt	ggaagcctgg	accgatgtaa	gggttcgggg	ctgtgccttt	tactgtctgt	3300
ggaagggggg	gggtgtctgc	cccaaaagca	gggcttcaat	taagaaatgc	ctctttgaaa	3360
gggtgtacct	gggtatccctg	tctgagggtg	actccagggt	gcgccacaat	gtggcctccg	3420
actgtgggtg	cttcatgcta	gtgaaaagcg	tggctgtgat	taagcataac	atggtagtg	3480
gcaactgcga	ggacagggcc	tctcagatgc	tgacctgtct	ggacggcaac	tgtcacctgc	3540
tgaagaccat	tcacgtagcc	agccactctc	gcaaggcctg	gccagtgttt	gagcataaca	3600
tactgaccgc	ctgttccttg	catttgggta	acaggagggg	gggtgtccta	ccttaccaat	3660
gcaatttgag	tcacactaag	atattgcttg	agcccgagag	catgtccaag	gtgaacctga	3720
acgggggtgt	tgacatgacc	atgaagatct	ggaagggtgct	gaggtacgat	gagaccgcga	3780
ccagggtgcag	accctgcgag	tgtggcggtg	aacatattag	gaaccagcct	gtgatgctgg	3840
atgtgaccga	ggagctgagg	cccgatcact	tgggtgctggc	ctgcaccgcg	gctgagtttg	3900
gctctagcga	tgaagataca	gattgaggtg	ctgaaatgtg	tgggcgtggc	ttaaggggtg	3960
gaaagaatat	ataagggtgg	ggctttatgt	agttttgtat	ctgttttgca	gcagccgcg	4020
ccgccatgag	caccaactcg	tttgatggaa	gcattgtgag	ctcatatttg	acaacgcgca	4080
tgccccatg	ggccgggggtg	cgtcagaatg	tgatgggctc	cagcattgat	ggtcgccccg	4140
tcttgcccgc	aaactctact	accttgacct	acgagaccgt	gtctggaacg	ccgttggaga	4200
ctgcagcctc	cgcgcgcgct	tcagccgctg	cagccaccgc	ccgcgggatt	gtgactgact	4260
ttgctttcct	gagcccgcct	gcaagcagtg	cagcttcccg	ttcatccgcc	cgcgatgaca	4320
agttgacggc	tcttttgcca	caattggatt	ctttgaccgc	ggaacttaat	gtcgtttctc	4380
agcagctggt	ggatctgcgc	cagcagggtt	ctgcccgtgaa	ggcttcctcc	cctcccaatg	4440
cgggtttaaaa	cataaataaa	aaaccagact	ctgtttggat	ttggatcaag	caagtgtctt	4500
gctgtctcag	ctgactgctt	aagtgcgaag	ccgaattgga	tccaattcgg	atcgatctta	4560
ttaaagcaga	acttgtttat	tgacgcttat	aatgggtaca	aataaagcaa	tagcatcaca	4620
aatttcacaa	ataaagcatt	tttttcaactg	cattctagtt	gtggtttgct	caaactcatc	4680
aatgtatctt	atcatgtctg	gtcgactcta	gactcttccg	cttcctcgct	cactgactcg	4740
ctgcgctcgg	tcgttcggct	gcggcgagcg	gtatcagctc	actcaaaggc	ggtaatacgg	4800
ttatccacag	aatcagggga	taacgcagga	aagaacatgt	gagcaaaaagg	ccagcaaaaag	4860
gccaggaacc	gtaaaaaggc	cgcgttgctg	gcgtttttcc	ataggctccg	ccccctgac	4920
gagcatcaca	aaaatcgacg	ctcaagtcag	aggtggcgaa	acccgacagg	actataaaga	4980
taccaggcgt	ttccccctgg	aagctccctc	gtgcgctctc	ctgttccgac	cctgccgctt	5040
accggatacc	tgtccgcctt	tctcccttcg	gttaagcgtgg	cgctttctca	tagctcacgc	5100
tgtagggtatc	tcagttcggt	gtaggtcggt	ggctccaagc	tgggctgtgt	gcacgaaccc	5160
cccgttcagc	ccgaccgctg	cgcccttatcc	ggtaactatc	gtcttgagtc	caaccgggta	5220
agacacgact	tatcgccact	ggcagcagcc	actggttaaca	ggattagcag	agcgagggtat	5280
gtaggcggtg	ctacagagtt	cttgaagtgg	tggcctaact	acggctacac	tagaaggaca	5340
gtatttggtg	tctgcgctct	gctgaagcca	gttaccttcg	gaaaaagagt	tggtagctct	5400
tgatccggca	aacaaaccac	cgctggtag	gggtggtttt	ttgtttgcaa	gcagcagatt	5460
acgcgcagaa	aaaaaggatc	tcaagaagat	cctttgatct	ttctacggg	gtctgacgct	5520
cagtggaaacg	aaaactcacg	ttaagggatt	ttggtcatga	gattatcaaa	aaggatcttc	5580
acctagatcc	ttttaaatga	aaaatgaagt	tttaaatcaa	tctaaagtat	atatgagtaa	5640
acttgggtctg	acagttacca	atgcttaatc	agtgaaggac	ctatctcagc	gatctgtcta	5700
tttcggtcat	ccatagttgc	ctgactcccc	gtcgtgtaga	taactacgat	acgggagggc	5760
ttaccatctg	gccccagtg	tgcaatgata	ccgcgagacc	cacgctcacc	ggctccagat	5820
ttatcagcaa	taaacagcgc	agccggaagg	gcccagcgca	gaagtgggtc	tgcaacttta	5880
tccgcctcca	tccagtctat	taattggtgc	cgggaagcta	gagtaagtag	ttcgccagtt	5940
aatagtttgc	gcaacgttgt	tgccattgct	acaggaatcg	tgggtgtcac	ctcgctgctt	6000
ggtaggtgct	cattcagctc	cggttcccaa	cgatcaaggc	gagttacatg	atcccccatg	6060
ttgtgcaaaa	aagcgggttag	ctccttcggt	cctccgatcg	ttgtcagaag	taagttggcc	6120
gcagtgttat	cactcatggt	tatggcagca	ctgcataaatt	ctcttactgt	catgccatcc	6180
gtaagatgct	tttctgtgac	tggtagtagc	tcaaccaagt	cattctgaga	atagtgtatg	6240
cggcgaccga	gttgctcttg	cccggcgctc	atacgggata	ataccgcgce	acatagcaga	6300
actttaaaag	tgctcatcat	tggaaaacgt	tcttcggggc	gaaaactctc	aaggatctta	6360

-11-

```

ccgctgttga gatccagttc gatgtaaccc actcgtgcac ccaactgatc ttcagcatct 6420
tttactttca ccagcgtttc tgggtgagca aaaacaggaa ggcaaaatgc cgcaaaaaag 6480
ggaataaggg cgacacggaa atgttgaata ctcatactct tcctttttca atattattga 6540
agcattttatc agggttattg tctcatgagc ggatacatat ttgaatgtat ttagaaaaat 6600
aaacaaatag gggttccgcg cacatttccc cgaaaagtgc cacctgacgt ctaagaaacc 6660
attattatca tgacattaac ctataaaaaat aggcgtatca cgaggcccct ttcgtctcgc 6720
gcgtttcggg gatgacgggtg aaaacctctg acacatgcag ctcccggaga cggtcacagc 6780
ttgtctgtaa gcggatgccg ggagcagaca agcccgtcag ggcgcgtcag cgggtgttgg 6840
cgggtgtcgg ggctggctta actatgcggc atcagagcag attgtactga gagtgcacca 6900
tatgcgggtg gaaataccgc acagatgcgt aaggagaaaa taccgcatca ggaaattgta 6960
agcgttaata ttttggttaa attcgcgtta aatttttgtt aaatcagctc attttttaac 7020
caataggccg aaatcggcaa aatcccttat aaatcaaaag aatagaccga gatagggttg 7080
agtgttgttc cagtttgtaa caagagtcga ctattaaaga acgtggactc caacgtcaaa 7140
gggcgaaaaa ccgtctatca gggcgatggc ccactacgtg aaccatcacc ctaatcaagt 7200
tttttggggg cgaggtgccg taaagcacta aatcggaaac cttaaaggag ccccgatttt 7260
agagcttgac ggggaaagcc ggcgaacgtg gcgagaaagg aagggaagaa agcgaaagga 7320
gcgggcgcta gggcgctggc aagtgtagcg gtcacgctgc gcgtaaccac cacaccgcgc 7380
gcgtttaatg cgccgctaca gggcgctgcc cattcgccat tcaggctgcg caactgttgg 7440
gaaggcgcat cgggtgcgggc ctcttcgcta ttacgccagc tggcgaaagg gggatgtgct 7500
gcaaggcgat taagttgggt aacgccaggg ttttccagc cagcagcttg taaaacgacg 7560
gccagtgaat tgtaatacga ctactatag ggcgaattaa ttcggggg 7607

```

<210> 42

<211> 11600

<212> DNA

<213> Artificial Sequence

<220>

<223> Plasmid MMTV-E2a-SV40-Neo

<400> 42

```

gaattccgca ttgcagagat attgtattta agtgcctagc tcgatacaat aaacgccatt 60
tgaccattca ccacattggt gtgcacctcc aagcttgggc agaaatgggt gaactcccga 120
gagtgtccta caccatgggg agaagcagcc aaggggttgt ttcccaccaa ggacgaccgc 180
tctgcgcaca aacggatgag cccatcagac aaagacatat tcattctctg ctgcaaaact 240
ggcatagctc tgctttgcct ggggctattg ggggaagtgt cggttcgtgc tcgcagggtc 300
ctcacccttg actcttttaa tagctcttct gtgcaagatt acaatctaaa caattcggag 360
aactcgacct tcctcctgag gcaaggacca cagccaactt cctcttacia gccgcctcga 420
ttttgtcctt cagaaaataga aataagaatg cttgctaaaa attatatatt taccaataag 480
accaatccaa taggttagat attagttact atgttaagaa atgaatcatt atcttttagt 540
actattttta ctcaaattca gaagttagaa atgggaatag aaaatagaaa gagacgctca 600
acctcaattg aagaacagggt gcaaggacta ttgaccacag gcctagaagt aaaaaaggga 660
aaaaagagtg tttttgtcaa aataggagac aggtgggtggc aaccagggac ttatagggga 720
ccttacatct acagaccaac agatgcccc ttaccatata caggaagata tgacttaaat 780
tgggataggt gggttacagt caatggctat aaagtgttat atagatccct cccttttcgt 840
gaaagactcg ccagagctag acctccttgg tgtatgttgt ctcaagaaga aaaagacgac 900
atgaaacaac aggtacatga ttatatattt cttaggaacag gaatgcactt ttggggaaag 960
attttccata ccaaggaggg gacagtggct ggactaatag aacattatct tgcaaaaact 1020
catggcatga gttattatga atagccttta ttggcccaac ctgcccgttc ccagggttta 1080
agtaagtttt tggttacaaa ctgttcttaa aacgaggatg tgagacaagt ggtttcctga 1140
cttgggttgg tatcaaagggt tctgatctga gctctgagtg ttctattttc ctatgttctt 1200
ttggaattta tccaaatctt atgtaaatgc ttatgtaaac caagatataa aagagtgtct 1260
atttttttag taaacttgca acagtccctc cattcacctc ttgtgtgttt gtgtctgttc 1320
gccatcccggt ctccgctcgt cacttatcct tcactttcca gaggggtccc ccgcagaccc 1380
cggcgaccct caggctcggcc gactgcggca gctggcgccc gaacagggac cctcggataa 1440
gtgacccttg totctatttc tactatttgg tgtttgtctt gtattgtctc tttcttgtct 1500
ggctatcctc acaagagcgg aacggactca ccatagggac caagctagcg cttctcgtcg 1560
cgtccaagac cctcaaagat ttttggcact tcggtgagcg aggcgatatt aggtatgaca 1620
gcgccttggc gcaaggccag ctgcttgtcc gctcggctgc ggttggcacg gcaggatagg 1680
ggatatcttg agttttggaa aaagatgtga taggtggcaa gcacctctgg cacggcaaat 1740
acggggtaga agttgaggcg cgggttgggc tcgcatgtgc cgttttcttg gcgtttgggg 1800

```

-12-

ggtacgcgcg	gtgagaatag	gtggcggttcg	taggcaaggc	tgacatccgc	tatggcgagg	1860
ggcacatcgc	tgcgctcttg	caacgcgtcg	cagataatgg	cgcaactggcg	ctgcagatgc	1920
ttcaacagca	cgctcgtctcc	cacatctagg	tagtcgccat	gcctttcgtc	ccccgcgccg	1980
acttgttcct	cgtttgccctc	tgcgttggtcc	tggctcttgct	ttttatcctc	tgttgggtact	2040
gagcgggtcct	cgctcgtcttc	gcttacaaaa	cctgggtcct	gctcgataat	cacttcctcc	2100
tcctcaagcg	ggggtgcctc	gacggggaag	gtggtaggcg	cgttggcggc	atcggtggag	2160
gcggtggtgg	cgaactcaga	gggggcgggt	aggctgtcct	tcttctcgac	tgactccatg	2220
atctttttct	gcctatagga	gaaggaaatg	gccagtcggg	aagaggagca	gcgcgaaacc	2280
acccccgagc	gcggacgcgg	tgcggcgcgga	cgccccccaa	ccatggagga	cggtgcgtcc	2340
ccgtccccgt	cgccgcgcgc	tccccgggcg	cccccaaaaa	agcggatgag	gcggcgtatc	2400
gagtccgagg	acgaggaaga	ctcatcacia	gacgcgtgg	tgccgcgcac	accagccccg	2460
cggccatcga	cctcggcggc	ggatttggtcc	attgcgccc	agaagaaaaa	gaagcgccct	2520
tctcccaagc	ccgagcgccc	gccatcacca	gaggtaatcg	tgacacgcga	ggaagaaaga	2580
gaagatgtgg	cgctacaaat	ggtgggtttc	agcaaccac	cggtgctaata	caagcatggc	2640
aaaggaggta	agcgcacagt	gcggcggtcg	aatgaagacg	acccagtggc	gcgtggtatg	2700
cggacgcgaag	aggaagagga	agagcccagc	gaagcggaaa	gtgaaattac	ggtgatgaac	2760
ccgctgagtg	tgccgatcgt	gtctgcgtgg	gagaagggca	tgagggctgc	gcgcgcgctg	2820
atggacaagt	accacgtgga	taacgatcta	aaggcgaact	tcaaactact	gcctgaccaa	2880
gtggaagctc	tggcggccgt	atgcaagacc	tggtgaacg	aggagcaccg	cggtgtgcag	2940
ctgaccttca	ccagcaaaaa	gacctttgtg	acgatgatgg	ggcgattcct	gcaggcgtag	3000
ctgcagtcgt	ttgcagaggt	gacctacaag	catcacgagc	ccacgggctg	cgcgttgtgg	3060
ctgcaccgct	gcgctgagat	cgaaggcgag	cttaagtgtc	tacacggaag	cattatgata	3120
aataaggagc	acgtgattga	aatggatgtg	acgagcgaaa	acgggcagcg	cgcgctgaag	3180
gagcagtcta	gcaaggccaa	gatcgtgaag	aacccggtggg	gccgaaatgt	ggtgcagatc	3240
tccaacaccg	acgcaaggtg	ctgcgtgcac	gacgcggcct	gtccggccaa	tcagttttcc	3300
ggcaagtctt	gcggcatgtt	cttctctgaa	ggcgcaaaag	ctcaggtggc	ttttaagcag	3360
atcaaggctt	ttatgcaggc	gctgtatcct	aacgcccaga	ccgggcacgg	tcaccttttg	3420
atgccactac	ggtgcgagtg	caactcaaag	cctgggcacg	cgcccttttt	gggaaggcag	3480
ctaccaaaag	tgactccgtt	cgccctgagc	aacgcggagg	acctggacgc	ggatctgatc	3540
tccgacaaag	cgctgctggc	cagcgtgcac	cacccggcgc	tgatagtgtt	ccagtgtgc	3600
aaccctgtgt	atcgcaactc	gcgcgcgcag	ggcggaggcc	ccaactgcga	cttcaagata	3660
tccggcggcg	acctgctaaa	cgcttgggtg	atggtgcgca	gcctgtggag	tgaaaacttc	3720
accgagctgc	cgcggtgggt	tgtgcctgag	tttaagtggg	gcactaaaca	ccagtatcgc	3780
aacgtgtccc	tgccagtggc	gcatagcgat	gcggcgagag	acccttttga	tttttaaacg	3840
gcgcagacgg	caagggtggg	ggtaataaat	cacccgagag	tgtacaaata	aaagcatttg	3900
cctttattga	aagtgtctct	agtacattat	ttttacatgt	ttttcaagtg	acaaaaagaa	3960
gtggcgctcc	taatctgcgc	actgtggctg	cggaagtagg	gcgagtggcg	ctccagggaag	4020
ctgtagagct	gttcctgggt	gcgacgcagg	gtgggtgtga	cctggggact	ggtgagcatg	4080
gagttgggta	ccccgtaat	aaggttcatg	tgggggttgt	gatccatggg	agtttggggc	4140
cagttggcaa	aggcgtggag	aaacatgcag	cagaatagtc	cacagcgccg	cgagttgggc	4200
ccctgtacgc	tttgggtgga	ctttccagc	gttatacagc	ggtcggggga	agaagcaatg	4260
gcgctacggc	gcaggagtga	ctcgtactca	aactggtaaa	cctgcttgag	tcgctgggtca	4320
gaaaagccaa	agggtcctaa	gaggtagcat	gtttttgagt	gcgggttcca	ggcaaaggcc	4380
atccagtgtg	cgcccccagt	ctcgcgaccg	gccgtattga	ctatggcgca	ggcgagcttg	4440
tgtggagaaa	caaagcctgg	aaagcgcttg	tcataaggtg	ccaaaaaata	tgccccacaa	4500
ccaagatctt	tgacaatggc	tttcagttcc	tgctcactgg	agcccatggc	ggcagctgtt	4560
gttgatgttg	cttgcttctt	tatgttgtgg	cgttgcggcg	cgagaagggc	gtgcgcagggt	4620
acacggtttc	gatgacgccg	cggtgcggcc	ggtgcacacg	gaccacgtca	aagacttcaa	4680
acaaaacata	aagaagggtg	ggctcgtcca	tgggatccat	atatagggcc	cggtttataa	4740
ttacctcagg	tcgacctcga	gggatctttg	tgaaggaaacc	ttacttctgt	ggtgtgacat	4800
aattggacaa	actacctaca	gagatttaaa	gctctaagggt	aaatataaaa	tttttaagtg	4860
tataatgtgt	taaactactg	attctaattg	tttgtgtatt	ttagattcca	acctatggaa	4920
ctgatgaatg	ggagcagttg	tggaatgcct	ttaatgagga	aaacctgttt	tgctcagaag	4980
aatgccatc	tagtgatgat	gaggctactg	ctgactctca	acattctact	cctccaaaaa	5040
agaagagaaa	ggtagaagac	cccaaggact	ttccttcaga	attgctaagt	tttttgagtc	5100
atgctgtgtt	tagtaataga	actcttgctt	gctttgctat	ttacaccaca	aaggaaaaag	5160
ctgcactgct	atacaagaaa	attatggaaa	aattctctgt	aacctttata	agtaggcata	5220
acagttataa	tcataacata	ctgttttttc	ttactccaca	caggcataga	gtgtctgcta	5280
ttaataacta	tgctcaaaaa	ttgtgtacct	ttagcttttt	aatttgtaaa	ggggttaata	5340
aggaatatct	gatgtatagt	gccttgacta	gagatcataa	tcagccatac	cacatttgta	5400
gagggttttac	ttgctttaaa	aaacctccca	cacctcccc	tgaacctgaa	acataaaatg	5460

-13-

aatgcaattg	ttgttggttaa	cttgttttatt	gcagcttata	atgggttacia	ataaagcaat	5520
agcatcacaa	atttcacaaa	ttaaagcattt	ttttcactgc	attctagtgtg	tggtttgtcc	5580
aaactcatca	atgtatctta	tcatgtcttg	atccggctgt	ggaatgtgtg	tcagttagg	5640
tgtggaaagt	ccccaggctc	cccagcaggc	agaagtatgc	aaagcatgca	tctcaattag	5700
tcagcaacca	gggtgtggaaa	gtccccaggc	tccccagcag	gcagaagtat	gcaaagcatg	5760
catctcaatt	agtcagcaac	catagtcccg	cccctaactc	cgcccatccc	gcccctaact	5820
cgcgccagtt	cgcgccattc	tccgccccat	ggctgactaa	ttttttttat	ttatgcagag	5880
gccgaggccg	cctcggcctc	tgagctatc	cagaagtatg	gaggaggctt	ttttggaggc	5940
ctaggctttt	gcaaaaagct	tcacgctgcc	gcaagcactc	agggcgcaag	ggctgctaaa	6000
ggaagcggaa	cacgtagaaa	gccagtccgc	agaaacggtg	ctgaccccg	atgaatgtca	6060
gctactgggc	tatctggaca	agggaaaacg	caagcgcaaa	gagaaagcag	gtagcttgca	6120
gtgggctttac	atggcgatag	ctagactggg	cggttttatg	gacagcaagc	gaaccggaat	6180
tgccagctgg	ggcgccctct	ggtaagggtg	ggaagccctg	caaagtaaac	tggatggctt	6240
tcttgccgcc	aaggatctga	tggcgagggg	gatcaagatc	tgatcaagag	acaggatgag	6300
gatcgtttctg	catgattgaa	caagatggat	tgacgcagag	ttctccggcc	gcttgggttg	6360
agaggctatt	cggctatgac	tgggcacaa	agacaatcgg	ctgctctgat	gccgcccgtg	6420
tccggctgtc	agcgcagggg	cgcccggttc	tttttgtcaa	gaccgacctg	tccggtgccc	6480
tgaatgaact	gcaggacgag	gcagcgcggc	tatcgtggct	ggccacgacg	ggcgttccct	6540
gcgagctgt	gctcgacgtt	gtcactgaag	cgggaaggga	ctggctgcta	ttgggcgaag	6600
tgccggggca	ggatctcctg	tcatctcacc	ttgctcctgc	cgagaaagta	tccatcatgg	6660
ctgatgcaat	cggtatgctg	catagcgttg	atccggctac	ctgcccattc	gaccaccaag	6720
cgaacatcg	catcgagcga	gcacgtactc	ggatggaagc	cggtcttgct	gatcaggatg	6780
atctggacga	agagcatcag	gggctcgcgc	cagccgaact	gttcgccagg	ctcaaggcgc	6840
gcatgccgca	cggcgaggat	ctcgtcgtga	cccattggcg	tgccctgctt	ccgaatatca	6900
tggtggaaaa	tgcccgcttt	tctggattca	tgcactgtgg	ccggctgggt	gtggcgacc	6960
gctatcagga	catagcgttg	gctaccgtg	atatgtctga	agagcttggc	ggcgaatggg	7020
ctgaccgctt	cctcgtgctt	tacggtatcg	ccgctcccga	ttcgcagcgc	atcgcttctt	7080
atcgcccttct	tgacgagttc	ttctgagcgg	gactctgggg	ttcgaaatga	ccgaccaagc	7140
gacgcccac	ctgccatcac	gagatttcga	ttccaccgcc	gccttctatg	aaaggttggg	7200
cttcggaate	gttttccggg	acgcggctg	gatgatcctc	cagcgccggg	atctcatgct	7260
ggagtctctc	gcccaccccg	ggctcgatcc	cctcgcgagt	tgggttcagct	gctgcctgag	7320
gctggacgac	ctcgcggagt	tctaccggca	gtgcaaatcc	gtcggcatcc	aggaaaccag	7380
cagcggctat	ccgcgcacat	atgccccga	actgcaggag	tggggaggca	cgatggccgc	7440
tttggtcccc	gatctttgtg	aaggaacctt	acttctgtgg	tgtgacataa	ttggacaaac	7500
tacctacaga	gatttaaagc	tctaaggtaa	atataaaatt	tttaagtgtg	taatgtgtta	7560
aactactgat	tctaattggt	tgtgtatttt	agattccaac	ctatggaact	gatgaatggg	7620
agcagtgggtg	gaatgccttt	aatgaggaaa	acctgttttg	ctcagaagaa	atgccatcta	7680
gtgatgatga	ggctactgct	gactctcaac	attctactcc	tccaaaaaag	aagagaaagg	7740
tagaagaccc	caaggacttt	ccttcagaat	tgttaagttt	tttgagtcat	gctgtgttta	7800
gtaatagaac	tcttgctatg	tttgctattt	acaccacaaa	ggaaaaagct	gcactgctat	7860
acaagaaaaat	tatggaaaaa	tattctgtaa	cctttataag	taggcataac	agttataatc	7920
ataacataact	gttttttctt	actccacaca	ggcatagagt	gtctgctatt	aataactatg	7980
ctcaaaaatt	gtgtaccttt	agctttttta	tttgtaaagg	ggttaataag	gaatatttga	8040
tgtatagtgc	cttgactaga	gatcataate	agccatacca	catttgtaga	ggttttactt	8100
gctttaaaaa	acctcccaca	cctccccctg	aacctgaaac	ataaaatgaa	tgcaattggt	8160
gttggttaact	tgttttattg	agcttataat	ggttacaaat	aaagcaatag	catcacaaat	8220
ttcacaaata	aagcattttt	ttcactgcat	tctagtgtgtg	gtttgtccaa	actcatcaat	8280
gtatctttatc	atgtctggat	ccccaggga	ctcctctgtg	tcctcataaa	ccctaaccctc	8340
ctctacttga	gaggacattc	caatcatagg	ctgcccattc	acctctgtgt	tcctcctggt	8400
aattaggtca	cttaacaaaa	aggaaattgg	gtaggggttt	ttcacagacc	gcttttctaag	8460
ggtaatttta	aaatatctgg	gaagtcctct	ccactgctgt	gttccagaag	tggttggttaa	8520
cagcccacaa	atgtcaacag	cagaaacata	caagctgtca	gctttgcaca	agggcccaac	8580
accctgctca	tcaagaagca	ctgtgggtgc	tgtgttagta	atgtgcaaaa	caggaggcac	8640
attttcccca	cctgtgtagg	ttccaaaata	tctagtgttt	tcattttttac	ttggatcagg	8700
aaccagcac	tccactggat	aagcattatc	cttatccaaa	acagccttgt	ggtcagtgtt	8760
catctgctga	ctgtcaactg	tagcattttt	tgggggttaca	gtttgagcag	gatatttggg	8820
cctgtagttt	gctaacacac	cctgcagctc	caaagggtcc	ccaccaacag	caaaaaaatg	8880
aaaatttgac	ccttgaatgg	gttttccagc	accattttca	tgagtttttt	gtgtccctga	8940
atgcaagttt	aacatagcag	ttaccccaat	aacctcagtt	ttaacagtaa	cagcttccca	9000
catcaaaaata	tttccacagg	tttaagtcctc	atttaaatta	ggcaaaggaa	ttcttgaaga	9060
cgaaggggcc	tcgtgatagc	cctattttta	taggttaatg	tcatgataat	aatggtttct	9120

-14-

```

tagacgtcag gtggcacttt tcggggaaat gtgcgcggaa cccctatttg tttatttttc 9180
taaatacatt caaatatgta tccgctcatg agacaataac cctgataaat gcttcaataa 9240
tattgaaaaa ggaagagtat gagtattcaa catttccgtg tcgcccttat tccctttttt 9300
gcggcatttt gccttcctgt ttttgctcac ccagaaacgc tgggtgaaagt aaaagatgct 9360
gaagatcagt tgggtgcacg agtgggttac atcgaactgg atctcaacag cggtaagatc 9420
cttgagagtt ttcgccccga agaacgtttt ccaatgatga gcacttttaa agttctgcta 9480
tgtggcgcggt tattatcccg tgttgacgcc gggcaagagc aactcggtcg ccgcatacac 9540
tattctcaga atgacttggg tgagtactca ccagtcacag aaaagcatct tacggatggc 9600
atgacagtaa gagaattatg cagtgcgtgcc ataaccatga gtgataaacac tgcggccaac 9660
ttacttctga caacgatcgg aggaccgaag gagctaaccg cttttttgca caacatgggg 9720
gatcatgtaa ctgcgcctga tcggtgggaa ccggagctga atgaagccat accaaacgac 9780
gagcgtgaca ccacgatgcc tgcagcaatg gcaacaacgt tgcgcaaact attaactggc 9840
gaactactta ctctagcttc ccggcaacaa ttaatagact ggatggaggc ggataaagtt 9900
gcaggaccac ttctgcgctc ggcccttccg gctggctggg ggccagatgg taagccctcc 10020
gccggtgagc gtgggtctcg cgggtatcatt gcagcactgg ggccagatgg taagccctcc 10020
cgtatcgtag ttatctacac gacggggagt caggcaacta tggatgaacg aaatagacag 10080
atcgctgaga tagtgccctc actgattaag cattggtaac tgcagacca agtttactca 10140
tatatacttt agattgattt ttttaattta aaaggatcta ggtgaagatc 10200
ctttttgata atctcatgac caaaatccct taacgtgagt tttcgttcca ctgagcgtca 10260
gaccccgtag aaaagatcaa aggatcttct tgagatcctt ttttctgcg cgtaatctgc 10320
tgcttgcaaa caaaaaaacc accgctacca gcggtgggtt gtttgccgga tcaagagcta 10380
ccaactcttt ttccgaaggt aactggcttc agcagagcgc agataccaaa tactgtcctt 10440
ctagtgtagc cgtagttagg ccaccacttc aagaactctg tagcaccgcc tacatacctc 10500
gctctgctaa tcctgttacc agtggtctgct gccagtggcg ataagtctgt tcttaccggg 10560
ttggactcaa gacgatagtt accggataag gcgcagcggg cgggctgaac ggggggttcg 10620
tgcacacagc ccagcttgga gcgaacgacc tacaccgaac tgagatacct acagcgtgag 10680
ctatgagaaa gcgccacgct tcccgaaggg agaaaggcgg acaggtatcc ggtaaagcggc 10740
agggctcgaa caggagagcg cacgaggagg ctccagggg gaaacgcctg gtatctttat 10800
agtcctgtcg ggtttcgcca cctctgactt gagcgtcgat ttttgtgatg ctcgtcaggg 10860
gggcgagacc tatggaaaaa cgccagcaac gcggcctttt tacggttcct ggcccttttg 10920
tggccttttg ctccatgtt ctttccctgc ttatcccctg attctgtgga taaccgtatt 10980
accgcctttg agtgagctga taccgctcgc cgacgcgaa cgaccgagcg cagcgagtca 11040
gtgagcgagg aagcggaaga gcgcctgatg cggatatttc tccttacgca tctgtgcggt 11100
atctcacacc gcatatggtg cactctcagt acaatctgct ctgatgccgc atagttaagc 11160
cagtatctgc tccctgcttg tgtgttgagg gtcgctgagt agtgcgcgag caaaatttaa 11220
gctacaacaa ggcaaggctt gaccgacaat tgcatgaaga atctgcttag ggtaggcgt 11280
tttgcgctgc ttgcgcatgt acgggcacaga tatacgcgta tctgagggga ctagggtgtg 11340
tttaggcgaa aagcggggct tcggttgtag gcggttagga gtccctcag gatatagtag 11400
tttcgctttt gcatagggag ggggaaatgt agtcttatgc aatacacttg tagtcttgca 11460
acatggtaac gatgagttag caacatgcct tacaaggaga gaaaaagcac cgtgcatgcc 11520
gattggtgga agtaaggtag tacgatcgtg ccttattagg aaggcaacag acgggtctga 11580
catggattgg acgaaccact

```

<210> 43

<211> 35211

<212> DNA

<213> Artificial Sequence

<220>

<223> Plasmid Av1nBg

<400> 43

```

catcatcaat aatatacctt attttggatt gaagccaata tgataatgag ggggtggagt 60
ttgtgacgtg gcgcggggcg tgggaacggg gcgggtgacg tagtagtgtg gcggaagtgt 120
gatgttgcaa gtgtggcgga acacatgtaa gcgacggatg tggcaaaagt gacgtttttg 180
gtgtgcgccc gtgtacacag gaagtgacaa ttttcgcgcg gttttaggcg gatgttgtag 240
taaatttggg cgtaaccgag taagatttgg ccattttcgc gggaaaactg aataagagga 300
agtgaaatct gaataatttt gtgttactca tagcgcgtaa tatttgtcta gggccgcggg 360
gactttgacc gtttacgttg agactcgcc aggtgttttt ctcaggtgtt ttccgcgttc 420
cgggtcaaag ttggcgttt attattatg tcagtacgta ccagtgcact ggccataggaa 480
gcttggtacc ggtgaattcg ctacgcttcg cgccccgatg tacgggcccag atatacgcgt 540

```


-15-

atctgagggg	actaggggtg	gttttaggcga	aaagcggggc	ttcggttgta	cgcggttagg	600
agtccctca	ggatatagta	gtttcgcttt	tgcataggga	gggggaaatg	tagtcttatg	660
caatactctt	gtagtcttgc	aacatggtaa	cgatgagtta	gcaacatgcc	ttacaaggag	720
agaaaaagca	ccgtgcatgc	cgattggtgg	aagtaagggtg	gtacgatcgt	gccttattag	780
gaaggcaaca	gacgggtctg	acatggattg	gacgaaccac	tgaattccgc	attgcagaga	840
tattgtatct	aagtgcctag	ctcgatacaa	taaacgccat	ttgaccattc	accacattgg	900
tgtgcaccte	cggccctggc	cactctcttc	cgcacgcgtg	tctgcggggg	ccagctgttg	960
ggctcgcggt	tgaggacaaa	ctcttcgcgg	tctttccagt	actcttggat	cgaaaacccg	1020
tcggcctccg	aacggtaact	cgccgcggag	ggacctgagc	gagtccgcat	cgaccggatc	1080
ggaaaacctc	tcgagaaagg	cgtgtaacca	gtcacagtcg	ctctagaact	agtggatccc	1140
ccgggctgca	ggaattcgat	ctagatggat	aaaggtccaa	aaaagaagag	aaaggtagaa	1200
gaccccaagg	actttccttc	agaattgcta	agttttttga	gtgattcact	ggccgtcggt	1260
ttacaacgtc	gtgactggga	aaacctggcg	gttacccaac	ttaatcgctt	tgacgacat	1320
ccccctttcg	ccagctggcg	taatagcgaa	gaggcccgca	ccgatcgccc	ttcccaacag	1380
ttgcgcagcc	tgaatggcga	atggcgcttt	gcctgggtttc	cggcaccaga	agcggtgccg	1440
gaaagctggc	tggagtgcga	tcttctctgag	gccgatactg	tcgtcgtccc	ctcaaactgg	1500
cgatgacacg	gttacgatgc	gcccactctac	accaaactgaa	cctatcccat	tacgggtcaat	1560
ccgcccgtttg	ttcccacgga	gaatccgacg	ggttgttact	cgctcacatt	taatgttgat	1620
gaaagctggc	tacaggaagg	ccagacgcga	attatTTTTg	atggcggttaa	ctcggcgttt	1680
catctgtggt	gcaacggggc	ctgggtcggt	tacggccagg	acagtcgttt	gccgtctgaa	1740
tttgacctga	gcgcattttt	acgcgcggga	gaaaaccgcc	tcgcggtgat	ggtgctgcgt	1800
tggagtgcag	gcagttatct	ggaagatcag	gatattgtgc	ggatgagcgg	cattttccgt	1860
gacgtctcgt	tgctgcataa	accgactaca	caaatacgcg	atttccatgt	tgccactcgc	1920
tttaatatg	atttcagccg	cgctgtactg	gaggctgaag	ttcagatgtg	cgccgagttg	1980
cgtagactacc	tacgggtaac	agtttcttta	tggcagggtg	aaacgcagggt	cgccagcggc	2040
accgcgcctt	tcggcggtga	aattatcgat	gagcgtgggtg	gttatgccga	tcgcgtcaca	2100
ctacgtctga	acgtcgaaaa	cccgaaaactg	tggagcgccg	aaatcccgaa	tctctatcgt	2160
gcgggtggtg	aactgcacac	cgccgacggc	acgctgattg	aagcagaagc	ctgcgatgtc	2220
ggtttccgcg	aggtgcggat	tgaaaatggt	ctgctgctgc	tgaacggcaa	gccgttgctg	2280
attcgaggcg	ttaaccgtca	cgagcatcat	cctctgcattg	gtcagggtcat	ggatgagcag	2340
acgatggtgc	aggatatcct	gctgatgaag	cagaacaact	ttaacgccgt	gcgctgttctg	2400
cattatccga	accatccgct	gtggtagacg	ctgtgcgacc	gctacggcct	gtatgtgggtg	2460
gatgaagcca	atattgaaac	ccacggcatg	gtgccaatga	atcgtctgac	cgatgatccg	2520
cgctggctac	cgccgatgag	cgaacgcgta	acgcgaatgg	tgacgcgcga	tcgtaatcac	2580
ccgagtgtga	tcattctggtc	gctgggggaat	gaatacaggc	acggcgctaa	tcacgacgcg	2640
ctgtatcgct	ggatcaaatc	tgtcgatcct	tcccgcggcg	tgacgatgta	aggcggcgga	2700
gccgacacca	ggcccaaccg	tattatttgc	ccgatgtacg	cgcgcggtgga	tgaagaccag	2760
cccttcccgg	ctgtgccgaa	atggtccatc	aaaaaatggc	tttcgctacc	tgagagagacg	2820
cgcccgcgtg	tcctttgcga	atacggccac	gcgatgggtg	acagtccttg	cggtttcgct	2880
aaatactggc	aggcggtttc	tcagtatccc	cgtttacagg	gcggcttcgt	ctgggactgg	2940
gtggatcagc	cgctgattaa	atatgatgaa	aacggcaacc	cgtggctcggc	ttacggcggtg	3000
gtatttggcg	atagcgcgaa	cgatcgccag	ttctgtatga	acggctcggg	ctttgcccgc	3060
cgcacgccgc	atccagcgct	gacggaagca	aaacaccagc	agcagttttt	ccagttccgt	3120
ttatccgggc	aaaccatcga	agtgaaccagc	gaatacctgt	tccgtcatag	cgataaccgag	3180
ctcctgcact	ggatggtggc	gctggatggt	aagccgctgg	caagcggtga	agtgcctctg	3240
gatgtcgctc	cacaaggtaa	acagttgatt	gaactgcctg	aactaccgca	gccggagagc	3300
gcggggcaac	tctggctcac	agtacgcgta	gtgcaaccga	acgcgaccgc	atgggtcagaa	3360
gccgggcaca	tcagcgctcg	gcagcagtg	cgtctggcgg	aaaacctcag	tgtgacgctc	3420
cccgcccgct	cccacgccat	cccgcatctg	accaccagcg	aaatggattt	ttgcatcgag	3480
ctgggttaata	agcgttggca	atttaaccgc	cagtcaggct	ttctttcaca	gatgtggatt	3540
ggcgataaaa	aacaactgct	gacgcccgtg	cgcgatcagt	tcacccgtgc	accgctggat	3600
aacgacattg	gcgtaagtga	agcgacccgc	attgacccta	acgcctgggt	cgaacgctgg	3660
aaggcgggcg	gccattacca	ggccgaagca	gcgttggtgc	agtgcacggc	agatacactt	3720
gctgatgcgg	tgctgattac	gaccgctcac	gcgtggcagc	atcaggggaa	aaccttattt	3780
atcagcggga	aaacctaccg	gattgatggt	agtggtcaaa	tggcgattac	cgttgatggt	3840
gaagtggcga	gcgatacacc	gcattccggc	cggattggcc	tgaactgccca	gctggcgag	3900
gtagcagagc	gggtaaactg	gctcggatta	gggcccgaag	aaaactatcc	cgaccgcctt	3960
actgccgcct	gttttgaccg	ctgggatctg	ccattgtcag	acatgtatac	cccgtacgtc	4020
ttcccagagc	aaaacgggtc	gcgctgcggg	acgcgcgaat	tgaattatgg	cccacaccag	4080
tggcgcgggc	acttccagtt	caacatcagc	cgtctcagtc	aacagcaact	gatggaaacc	4140
agccatcgcc	atctgctgca	cgcggaagaa	ggcacatggc	tgaatatcga	cggtttccat	4200

-16-

atgggggattg	gtggcgacga	ctcctggagc	ccgtcagtat	cggcggaatt	tcagctgagc	4260
gccggtcgct	accattacca	gttgggtctgg	tgtcaaaaat	aataatctcg	aatcaagctt	4320
atcgataccg	tcgaaacttg	tttattgcag	cttataatgg	ttacaaataa	agcaatagca	4380
tcacaaaattt	cacaaaataaa	gcattttttt	cactgcattc	tagttgtggg	ttgtccaaac	4440
tcacaaatgt	atcttatcat	gtctggatcc	gacctcggat	ctggaagggtg	ctgaggtacg	4500
atgagacccg	caccaggtgc	agaccctgcg	agtgtggcgg	taaacatatt	aggaaccagc	4560
ctgtgatgct	ggatgtgacc	gaggagctga	ggcccgatca	cttgggtgctg	gcctgcaccc	4620
gcgtgaggt	tggctctagc	gatgaagata	cagattgagg	tactgaaatg	tgtgggcgtg	4680
gcttaagggt	gggaaagaat	atataagggtg	ggggctttat	gtagttttgt	atctgttttg	4740
cagcagccgc	cgccgccatg	agcaccaact	cgtttgatgg	aagcattgtg	agctcatatt	4800
tgacaacgcg	catgccccca	tggggccggg	tgcgtcagaa	tgtgatgggc	tccagcattg	4860
atggtcgccc	cgctcctgccc	gcaaactcta	ctaccttgac	ctacgagacc	gtgtctggaa	4920
cgccgttgga	gactgcagcc	tccgccgcgc	cttcagccgc	tgcagccacc	gcccgcggga	4980
ttgtgactga	ctttgctttc	ctgagcccg	ttgcaagcag	tgcagcttcc	cgttcatccg	5040
cccgcgatga	caagttgacg	gctcttttgg	cacaattgga	ttctttgacc	cggaactta	5100
atgtcgtttc	tcagcagctg	ttggatctgc	gccagcaggt	ttctgccctg	aaggcttcc	5160
cccctcccaa	tgcgggttaa	aacataaata	aaaaaccaga	ctctgttttg	atttggtatca	5220
agcaagtgtc	ttgctgtcct	tatttagggg	ttttgcgcgc	gcggtaggcc	cgggaccagc	5280
ggtctcggtc	gttgagggtc	ctgtgtat	tttccaggac	gtggtaaagg	tgactctgga	5340
tgttcagata	catgggcata	agcccgtctc	tggggtggag	gtagcaccac	tgcagagctt	5400
catgctgcgg	ggtgggtgtt	tagatgatcc	agtcgtagca	ggagcgtcgg	gcgtgggtgc	5460
taaaaatgtc	tttcagtagc	aagctgattg	ccaggggcag	gcccttggtg	taagtgttta	5520
caaagcgggt	aagctgggat	gggtgcatac	gtggggatat	gagatgcac	ttggactgta	5580
tttttaggtt	ggctatgttc	ccagccatat	ccctccgggg	attcatgttg	tgcagaacca	5640
ccagcacagt	gtatccggtg	cacttgggaa	at ttgtcatg	tagcttagaa	ggaaatgcgt	5700
ggaagaactt	ggagacgccc	ttgtgacctc	caagat tttc	catgcattcg	tccataatga	5760
tggaatggg	cccacgggcg	gcggcctggg	cgaagatatt	tctgggatca	ctaacgtcat	5820
agtgtgttcc	aggatgagat	cgctataggc	catttttaca	aagcgcgggc	ggagggtgcc	5880
agactgcggt	ataatgggtc	catccggccc	aggggcgtag	ttaccctcac	agattttgcat	5940
ttcccacgct	ttgagttcag	atggggggat	catgtctacc	tgcggggcga	tgaagaaaac	6000
ggtttccggg	gtaggggaga	tcagctggga	agaaagcagg	ttcctgagca	gctgcgactt	6060
accgcagccg	gtgggcccgt	aaatcacacc	tattaccggg	tgcaactggt	agttaaagaa	6120
gctgcagctg	ccgtcatccc	tgagcagggg	ggccacttcg	ttaagcatgt	ccctgactcg	6180
catgttttcc	ctgaccaa	ccgcagaaag	gcgctcgccg	cccagcgata	gcagttcttg	6240
caaggaagca	aagtttttca	acggtttgag	accgtccgcc	gtaggcatgc	ttttgagcgt	6300
ttgacaaagc	agttccaggc	ggcccacag	ctcggtcacc	tgctctacgg	catctcgatc	6360
cagcatatct	cctcgtttcg	cgggttgggg	cggctttcgc	tgtacggcag	tagtcggtgc	6420
tcgtccagac	gggcccagggt	catgtctttc	cacgggcgca	gggtcctcgt	cagcgtagtc	6480
tgggtcacgg	tgaaggggtg	cgctccgggc	tgcgcgctgg	ccagggtgcg	cttgaggctg	6540
gtcctgctgg	tgctgaagcg	ctgccggtct	tgcgcctgcg	cgtcggccag	gtagcatttg	6600
accatggtgt	catagtccag	cccctccgcg	gcgtggccct	tggcgcgcag	cttgcccttg	6660
gaggagcgcg	cgcacgaggg	gcagtgcaga	cttttgaggg	cgtagagctt	gggcgcgaga	6720
aataccgatt	ccggggagta	ggcatccgcg	ccgcaggccc	cgcagacggt	ctcgcattcc	6780
acgagccagg	tgagctctgg	ccgttcgggg	tcaaaaacca	gggttccccc	atgctttttg	6840
atgcgtttct	tacctctggt	ttccatgagc	cggtgtccac	gctcggtgac	gaaaaggctg	6900
tccgtgtccc	cgtatacaga	cttgagaggc	ctgtcctcga	gcggtgttcc	gcggtcctcc	6960
tcgtatagaa	actcggacca	ctctgagaca	aaggctcgcg	tccaggccag	cacgaaggag	7020
gctaagtggg	aggggtagcg	gtcgttgtcc	actagggggg	ccactcgctc	cagggtgtga	7080
agacacatgt	cgccctcttc	ggcatcaagg	aagggtgattg	gtttgtaggt	gtaggccacg	7140
tgacacgggtg	ttcctgaagg	ggggctataa	aaggggggtg	gggcgcgctc	gtcctcactc	7200
tcttccgcat	cgctgtctgc	gagggccagc	tgttgggggtg	agtactccct	ctgaaaagcg	7260
ggcatgactt	ctgcgctaag	attgtcagtt	tccaaaaacg	aggaggattt	gatattcacc	7320
tggcccgcgg	tgatgccttt	gaggggtggc	gcattccatct	ggtcagaaaa	gacaatcttt	7380
ttgttgtcaa	gcttggtggc	aaacgacccg	tagagggcgt	tggacagcaa	cttggcgatg	7440
gagcgcaggg	tttggttttt	gtcgcgactc	gcgcgctcct	tggccgcgat	gttttagctg	7500
acgtattcgc	gcgcaacgca	ccgccattcg	ggaaagacgg	tgggtgcgctc	gtcgggcacc	7560
aggtgcacgc	gccaaccgcg	gttgtgcagg	gtgacaagggt	caacgctggt	ggctacctct	7620
ccgcgtaggc	gctcgttggt	ccagcagagg	cggccgcctc	tgcgcgagca	gaatggcggt	7680
aggggggtcta	gctcgcgtctc	gtccgggggg	tctgcgtcca	cggtaaagac	cccgggcagc	7740
agggcgcgct	cgaagtagtc	tatcttggat	ccttgcaagt	ctagcgcctg	ctgccatgcg	7800
cgggcggcaa	gcgcgcgctc	gtatgggttg	agtgggggac	cccatggcat	gggggtgggtg	7860

-17-

agcgcggagg	cgtacatgcc	gcaaatgtcg	taaacgtaga	ggggctctct	gagtatcca	7920
agatatgtag	ggtagcatct	tccaccgcgg	atgctggcgc	gcacgtaatc	gtatagttcg	7980
tgcgagggag	cgaggaggtc	gggaccgagg	ttgctacggg	cgggctgctc	tgctcggaag	8040
actatctgcc	tgaagatggc	atgtgagttg	gatgatatgg	ttggacgctg	gaagacgttg	8100
aagctggcgt	ctgtgagacc	taccgcgtca	cgcacgaagg	aggcgtagga	gtcgcgcagc	8160
ttgttgacca	gctcggcggt	gacctgcacg	tctagggcgc	agtagtccag	ggtttccttg	8220
atgatgtcat	acttatcctg	tccctttttt	ttccacagct	cgcggttgag	gacaaaactct	8280
tcgcggtctt	tccagtactc	ttggatcgga	aaocccgtcg	cctccgaacg	gtaagagcct	8340
agcatgtaga	actggttgac	ggcctggtag	gcgacgcatc	ccttttctac	gggtagcgcg	8400
tatgcctgcg	cggccttccg	gagcagaggtg	tgggtgagcg	caaaggtgtc	cctgaccatg	8460
actttgaggt	actggtattt	gaagtcagtg	tcgtcgcac	cgcctgctc	ccagagcaaa	8520
aagtcctgtc	gcttttttga	acgcggattt	ggcagggcga	aggtgacatc	gttgaagagt	8580
atctttcccg	cgcgagggcat	aaagttgcgt	gtgatgcgga	agggctcccg	cacctcgga	8640
cggttgttaa	ttacctgggc	ggcgagcacg	atctcgtcaa	agccgttgat	gttgtggccc	8700
acaatgtaaa	gttccaagaa	gcgcgggatg	cccttgatgg	aaggcaattt	tttaagttcc	8760
tcgtaggtga	gctcttcagg	ggagctgagc	ccgtgctctg	aaagggccca	gtctgcaaga	8820
tgagggttgg	aagcgacgaa	tgagctccac	aggtcacggg	ccattagcat	ttgcaggtgg	8880
tcgcgaaagg	tcctaaactg	gcgacctatg	gccatttttt	ctggggtgat	gcagtagaag	8940
gtaagcgggt	cttggtccca	gcggtcccat	ccaaggttcg	cggctaggtc	tcgcgcggca	9000
gtcactagag	gctcatctcc	gccgaacttc	atgaccagca	tgaagggcac	gagctgcttc	9060
ccaaaggccc	ccatccaagt	ataggtctct	acatcgtagg	tgacaaagag	acgctcggtg	9120
cgaggatgcg	agccgatcgg	gaagaactgg	atctcccgcc	accaattgga	ggagtggcta	9180
ttgatgtggt	gaaagtagaa	gtccctgcga	cgggcccgaac	actcgtgctg	gcttttgtaa	9240
aaacgtgcgc	agtactggca	gcggtgcacg	ggctgtacat	cctgcacgag	gttgacctga	9300
cgaccgcgca	caaggaagca	gagtgggaat	ttgagccctt	cgcctggcgg	gtttggctgg	9360
tggtcttcta	cttcggctgc	ttgtccttga	ccgtctggct	gctcgagggg	agttacggtg	9420
gatcggaacca	ccacgcgcgc	cgagcccaaa	gtccagatgt	ccgcgcgcgc	cggctcgagc	9480
ttgatgacaa	catcgcgag	atgggagctg	tccatggctc	ggagctccc	cggcgtcagg	9540
tcaggcggga	gctcctgcag	gtttacctcg	catagacggg	tcagggcgcg	ggctagatcc	9600
aggtgatacc	taattttccag	gggctgggtg	gtggcggcgt	cgatggcttg	caagaggccg	9660
catccccgcg	gcgcgactac	ggtaccgcgc	ggcgggcggt	gggcccgcgg	ggtgtccttg	9720
gatgatgcac	ctaaaagcgg	tgacgcgggc	gagcccccg	aggtaggggg	ggctccggac	9780
ccgcccgggag	agggggcagg	ggcacgtcgg	cgccgcgcgc	gggcaggagc	tggtgctgcg	9840
cgcgtaggtt	gctggcgaa	gcgacgacgc	ggcggttgat	ctcctgaatc	tggcgcctct	9900
gcgtgaagac	gacggggccc	gtgagcttga	gcctgaaaga	gagttcgaca	gaatcaattt	9960
cgggttcggt	gacggcgccc	tggcgcaaaa	tctcctgcac	gtctcctgag	ttgtcttgat	10020
aggcgatctc	ggccatgaac	tgctcgatct	cttccctctg	gagatctccg	cgctccggctc	10080
gctccacggg	ggcggcgagg	tcgttgga	tgccggccat	gagctgcgag	aaggcggtga	10140
ggcctccctc	gttccagacg	cggctgtaga	ccacgcccc	ttcgcatcgc	cgggcgcgca	10200
tgaccacctg	cgcgagattg	agctccacgt	ggcgggcgaa	gacggcgtag	tttcgacggc	10260
gtgaaaagag	gtagttcgag	gtggtggcgg	tgtgttctgc	cacgaagaag	tacataaccc	10320
agcgtcgcaa	cgtggattcg	ttgatatccc	ccaaggcctc	aaggcgctcc	atggcctcgt	10380
agaagtccac	ggcgaagtgg	aaaaactggg	agttgcgcgc	cgacacggtt	aactcctcct	10440
ccagaagacg	gatgagctcg	gcgacagtgt	cgcgcacctc	gcgctcaaag	gctacagggg	10500
cctcttcttc	ttcttcaatc	tcctcttcca	taagggcctc	cccttcttct	tcttctggcg	10560
gcggtggggg	agggggggaca	cggcggcgac	gacggcgcac	cgggaggcgg	tcgacaaagc	10620
gctcgatcat	ctccccgcgg	cgacggcgca	tggtctcggt	gacggcgcg	ccgttctcgc	10680
gggggcgcag	ttggaagacg	ccgcccgtca	tgtcccgggt	atgggttggc	ggggggctgc	10740
catgcccagc	ggatacggcg	ctaacgatgc	atctcaacaa	ttgttggtga	ggtactccgc	10800
cgcgagggga	cctgagcgag	tccgcatcga	ccggatcgga	aaacctctcg	agaaaggcgt	10860
ctaaccagtc	acagtcgcaa	ggtaggctga	gcaccgtggc	gggcggcgag	gggcggcggt	10920
cggggttggt	tctggcggag	gtgctgctga	tgatgtaatt	aaagtaggcg	gtcttgagac	10980
ggcggtatgg	cgacagaagc	accatgtcct	tgggtccggc	ctgctgaatg	cgcaggcggt	11040
cggccatgac	ccaggcttcg	ttttgacatc	ggcgaggctc	tttgtagtag	tcttgcatga	11100
gcctttcttc	cggcacttct	tcttctcctt	cctcttgctc	tgcatctctt	gcatctatcg	11160
ctgcggcggc	ggcggagttt	ggccgtagggt	ggcgccctct	tcctcccatg	cgtgtgaccc	11220
cgaagccctc	catcggctga	agcagggcta	ggtcggcgac	aacgcgctcg	gctaataatg	11280
cctgctgcac	ctgcgtgagg	gtagactgga	agtcacccat	gtccacaaa	cgggtggtag	11340
cgcctgtggt	gatggtgtaa	gtgcagttgg	ccataacgga	ccagttaacg	gtctggtgag	11400
ccggctgcga	gagctcgggtg	tacctgagac	gcgagtaagc	cctcgagtca	aatacgtagt	11460
cgttgcaagt	ccgcaccagg	tactggtatc	ccacccaaaa	gtgcggcggc	ggctggcggt	11520

-18-

agagggggcca	gcgtaggggtg	gccgggggctc	cggggggcgag	atctttccaac	ataaggcgat	11580
gatatccgta	gatgtacctg	gacatccagg	tgatgccggc	ggcgggtggtg	gaggcgcgcg	11640
gaaagtgcgc	gacgcgggttc	cagatgttgc	gcagcggcaa	aaagtgtccc	atggtcggga	11700
cgctctggcc	ggtcaggcgc	gcgcaatcgt	tgacgctcta	gaccgtgcaa	aaggagagcc	11760
tgtaagcggg	cactcttccg	tggtctgggtg	gataaattcg	caagggatc	atggcggacg	11820
accgggggttc	gagccccgta	tccggccgctc	cgccgtgatc	catgcggtta	ccgcccgcgt	11880
gtcgaaccca	ggtgtgcgac	gtcagacaac	gggggagtg	tccttttggc	ttccttccag	11940
gcgcggcgcc	tgctgcgcta	gcttttttgg	ccactggccg	cgcgacgcgt	aagcgggttag	12000
gctggaaagc	gaaagcatta	agtggctcgc	tcctgttagc	cggagggtta	ttttccaagg	12060
ggtgagtcgc	gggacccccg	gttcgagctc	cggaccggcc	ggactgcggc	gaacgggggtg	12120
ttgcctcccc	gtcatgcaag	accccgtttg	caaattcctc	cggaaacagg	gacgagcccc	12180
ttttttgctt	ttcccagatg	catccgggtg	tgccgcagat	gcgccccctc	cctcagcagc	12240
ggcaagagca	agagcagcgg	cagacatgca	gggcaccctc	ccctcctcct	accgcgtcag	12300
gagggggcgac	atccgcgggtt	gacgcggcag	cagatgggtga	ttacgaaccc	ccgcggcgcc	12360
gggcccggca	ctacctggac	ttggaggagg	gcgagggcct	ggcgcggtta	ggagcgccct	12420
ctcctgagcg	gtacccaagg	gtgcagctga	agcgtgatac	gcgtgaggcg	tacgtgccgc	12480
ggcagaacct	gtttcgcgac	cgcgaggagg	aggagcccga	ggagatgcgg	gatcgaaagt	12540
tcacgcagag	gcgcgagctg	cggcatggcc	tgaatcgca	gcggttgctg	cgcgaggagg	12600
actttgagcc	cgacgcgcga	accgggatta	gtcccgcgcg	cgcacacgtg	gcggccgcgc	12660
acctggtaac	cgcatacgag	cagacgggtga	accaggagat	taactttcaa	aaaagcttta	12720
acaaccacgt	gcgtacgctt	gtggcgcgcg	aggaggtggc	tataggactg	atgcatctgt	12780
gggactttgt	aagcgcgctg	gagcaaaacc	caaatagcaa	gccgctcatg	gcgcagctgt	12840
tccttatagt	gcagcacagc	agggacaacg	aggcattcag	ggatgcgctg	ctaaacatag	12900
tagagcccg	gggcccgtgg	ctgctcgatt	tgataaacat	cctgcagagc	atagtgggtg	12960
aggagcgcag	cttgagcctg	gctgacaagg	tggccgccat	caactattcc	atgcttagcc	13020
tgggcaagtt	ttacgcccgc	aagatatacc	atacccttta	cgttcccata	gacaaggagg	13080
taaagatcga	gggggttctac	atgcgcatgg	cgctgaaggt	gcttaccttg	agcgacgacc	13140
tgggcgttta	tcgcaacgag	cgcatccaca	aggccgtgag	cgtagcccg	cggcgcgagc	13200
tcagcgaccg	cgagctgatg	cacagcctgc	aaagggccct	ggctggcacg	ggcagcgcg	13260
atagagaggc	cgagtcctac	tttgacgcgg	gcgctgacct	gcgctgggcc	ccaagccgac	13320
gcgcccctga	ggcagctggg	gccggacctg	ggctggcggt	ggcaccgcgc	cgcgctggca	13380
acgtcgccgg	cgtggaggaa	tatgacgagg	acgatgagta	cgagccagag	gacggcgagt	13440
actaagcggg	gatgtttctg	atcagatgat	gcgaagcgca	acggaccgcg	cggtgcccgg	13500
ggcgctgcag	agccagccgt	ccggccttaa	ctccacggac	gactggcgcc	aggctcatgga	13560
ccgcatcatg	tcgctgactg	cgcgcaatcc	tgacgcgttc	cggcagcagc	cgcaggccaa	13620
ccggctctcc	gcaattcttg	aagcgggtgg	cccggcgcg	gcaaacccca	cgcacgagaa	13680
ggctgctggc	atcgtaaacg	cgctggccga	aaacagggcc	atccggcccg	acgagggcgg	13740
cctgggtctac	gacgcgctgc	ttcagcgctg	ggctcggtac	aacagcgcca	acgtgcagac	13800
caacctggac	cggctgggtg	gggatgtgcg	cgaggccgtg	gcgcagcgtg	agcgcgcgca	13860
gcagcagggc	aacctgggct	ccatgggtgc	actaaacgcc	ttcctgagta	cacagcccgc	13920
caacgtgcgc	cggggacagg	aggactacac	caacttttgt	agcgactgc	ggctaattggt	13980
gactgagaca	ccgcaaaagt	agggtgtacca	gtctggggcca	gactattttt	tccagaccag	14040
tagacaaggc	ctgcagaccg	taaacctgag	ccaggctttc	aaaaacttgc	aggggctgtg	14100
gggggtgcgg	gctcccacag	gcgaccgcgc	gaccgtgtct	agcttgctga	cgcccaactc	14160
gcgcctgttg	ctgctgctaa	tagcgccctt	cacggacagt	ggcagcgtgt	cccggggacac	14220
atacctaggt	cacttgctga	cactgtaccg	cgaggccata	ggtcaggcgc	atgtggacga	14280
gcatactttc	caggagatta	caagtgtcag	ccgcgcgctg	gggcaggagg	acacgggcag	14340
cctggaggca	accctaaact	acctgctgac	caaccggcgg	cagaagatcc	cctcgttgca	14400
cagttttaaac	agcgaggagg	agcgcatttt	gcgctacgtg	cagcagagcg	tgagccttaa	14460
cctgatgcgc	gacggggtaa	cgcccagcgt	ggcgctggag	atgaccgcgc	gcaacatgga	14520
accggggcatg	tatgcctcaa	accggccgtt	tatcaaccgc	ctaattggact	acttgcatcg	14580
cgcgggccg	gtgaaccccg	agtattttcac	caatgccatc	ttgaacccgc	actggctacc	14640
gccccctggt	ttctacaccg	ggggattcga	ggtgcccag	ggtaacgatg	gattcctctg	14700
ggacgacata	gacgacagcg	tggtttcccc	gcaaccgcag	accctgctag	agttgcaaca	14760
gcgcgagcag	gcagaggcgg	cgctgcgaaa	ggaaagcttc	cgcaggccaa	gcagcttgct	14820
cgatctaggc	gctgcggccc	cgcggtcaga	tgctagtagc	ccattttccaa	gcttgatagg	14880
gtctcttacc	agcactcgca	ccaccgcgcc	gcgcctgctg	ggcgaggagg	agtacctaaa	14940
caactcgctg	ctgcagccgc	agcgcgaaaa	aaacctgcct	ccggcatttc	ccaacaacgg	15000
gatagagagc	ctagtggaca	agatgagtag	atggaagacg	tacgcgcagg	agcacaggga	15060
cgtagccagc	ccgcggccgc	ccaccgctcg	tcaaaggcac	gaccgtcagc	gggggtctggt	15120
gtgggaggac	gatgactcgg	cagacgacag	cagcgtcctg	gatttgggag	ggagtggcaa	15180

cccgtttgcg	caccttcgcc	ccaggctggg	gagaatgttt	taaaaaaaaa	aaagcatgat	15240
gcaaaataaaa	aaactcacca	aggccatggc	accgagcggt	ggttttcttg	tattcccctt	15300
agtatgcggc	gcgcggcgat	gtatgaggaa	ggctcctctc	cctcctacga	gagtgtgggtg	15360
agcgcggcgc	cagtggcggc	ggcgctgggt	tctcccttcg	atgctcccct	ggacccgcgc	15420
tttgtgcctc	cgcggtacct	ggggcctacc	ggggggagaa	acagcatccg	ttactctgag	15480
ttggcacccc	tattcgacac	caccgcgtgtg	tacctgggtg	acaacaagtc	aacggatgtg	15540
gcataccctga	actaccagaa	cgaccacagc	aactttctga	ccacggtcac	tcaaaacaat	15600
gactacagcc	cgggggaggc	aagcacacag	accatcaatc	ttgacgaccg	gtcgcactgg	15660
ggcggcgacc	tgaaaacccat	cctgcatacc	aacatgccaa	atgtgaacga	gttcatgttt	15720
accaataagt	ttaaggcgcg	ggtgatgggt	tcgcgcttgc	ctactaagga	caatcagggtg	15780
gagctgaaat	acgagtgggt	ggagttcacg	ctgcccggag	gcaactactc	cgagaccatg	15840
accatagacc	ttatgaacaa	cgcgatcggt	gagcactact	tgaagtgagg	cagacagaac	15900
gggggttctg	aaagcgacat	cggggtaaag	tttgacaccc	gcaacttcag	actgggggtt	15960
gaccccgctc	ctgggtcttgt	catgcctggg	gtatatacaa	acgaagcctt	ccatccagac	16020
atcattttgc	tgccaggatg	cggggtggag	ttcacccaca	gccgcctgag	caacttgttg	16080
ggcatccgca	agcggcaacc	cttcaggag	ggcttttaga	tcacctacga	tgatctggag	16140
ggtggttaaca	ttcccgcact	gttggatgtg	gagcgctacc	aggcgagctt	gaaagatgac	16200
accgaacagg	gcgggggttg	cgcaggcggc	agcaacagca	gtggcagcgg	cgcggaagag	16260
aactccaacg	cggcagccgc	ggcaatgcag	ccggtggagg	acatgaacga	tcatgccatt	16320
cgcgcgacga	cctttgccac	acgggctgag	gagaagcgcg	ctgaggccga	agcagcggcc	16380
gaagctggcg	ccccgcctgc	gcaaccgcag	gtcgagaagc	ctcagaagaa	accggtgatt	16440
aaaccctga	cagaggacag	caagaaacgc	agttacaacc	taataagcaa	tgacagcacc	16500
ttcacccagt	accgcagctg	gtaccttgca	tacaactacg	gcgaccctca	gaccggaatc	16560
cgctcatgga	ccctgctttg	cactcctgac	gtaacctgcy	gctcggagca	ggtctactgg	16620
tcgattgccc	acatgatgca	agaccccgtg	accttccgct	ccacgcgcca	gatcagcaac	16680
tttccgggtg	tgggcgcgca	gctgtttgcc	gctgactcca	agagcttcta	caacgaccag	16740
gccgtctact	cccaactcat	ccgccagttt	acctctctga	cccacgtgtt	caatcgcttt	16800
cccgagaacc	agattttggc	gcgcccgcga	gccccacca	tcaccaccgt	cagtgaaaac	16860
gttcctgctc	tcacagatca	cgggacgcta	ccgctgcgca	acagcatcgg	aggagtccag	16920
cgagtgacca	ttactgacgc	cagacgcgcg	acctgcccct	acgtttacaa	ggccctgggc	16980
atagtctcgc	cgcgcgctcct	atcgagccgc	actttttgag	caagcatgtc	catccttata	17040
tcgcccagca	ataacacagg	ctggggcctg	cgcttcccaa	gcaagatgtt	tggcggggcc	17100
aagaagcgct	ccgaccaaca	cccagtgcgc	gtgcgcgggc	actaccgcgc	gccctggggc	17160
gcgcacaaac	gcggccgcac	tgggcgcacc	accgtcgatg	acgccatcga	cgcggtgggtg	17220
gaggaggcgc	gcaactacac	gcccacgcgc	ccaccagtgt	ccacagtgga	cgcgccatt	17280
cagaccgtgg	tgcgcgggagc	ccggcgctat	gctaaaatga	agagacggcg	gaggcgcgta	17340
gcacgtcgcc	accgcgcgcg	accgggact	gccgcccac	gcgcggcggc	ggccctgctt	17400
aaccgcgcac	gtcgcacccg	ccgacggggc	gccatgcggg	ccgctcgaag	gctggccgcg	17460
ggatttgtca	ctgtgcccc	caggtccagg	cacgagcggg	ccgccgcagc	agccgcggcc	17520
attagtgtca	tgactcaggg	tcgcaggggc	acagtgattt	gggtgcgcga	ctcggttagc	17580
ggcctgcgcg	tgcccgtgcg	caccgcgcgc	ccgcgcaact	agattgcaag	aaaaaactac	17640
ttagactcgt	actgttgtat	gtatccagcg	gcggcgcgcg	gcaacgaagc	tatgtccaag	17700
cgcaaaatca	agaagagat	gctccaggtc	atcgcgccgg	agatctatgg	ccccccgaag	17760
aaggaagagc	aggattacaa	gccccgaaa	ctaaagcggg	tcaaaaagaa	aaagaaagat	17820
gatgatgatg	aacttgacga	cgaggtggaa	ctgctgcacg	ctaccgcgcc	caggcgacgg	17880
gtacagtggg	aaggtcgacg	cgtaaaacgt	gttttgcgac	ccggcaccac	cgtagtcttt	17940
acgcccgggtg	agcgctccac	ccgcacctac	aagcgcggtg	atgatgaggt	gtacggcgac	18000
gaggacatgc	ttgagcaggc	caacgagcgc	ctcggggagt	ttgcctacgg	aaagcggcac	18060
aaggacatgc	tggcgttgcc	gctggacgag	ggcaacccaa	cacctagcct	aaagcccgta	18120
acactgcagc	agggtgctgc	cgcgcttgca	ccgtccgaag	aaaagcgcgg	cctaaagcgc	18180
gagtcctggtg	acttggcacc	caccgtgcag	ctgatggtag	ccaagcgcca	gcgactggaa	18240
gatgtcttgg	aaaaaatgac	cgtggaacct	gggctggagc	ccgaggtccg	cgtgcggcca	18300
atcaagcagc	tggcgccggg	actgggcgtg	cagaccgtgg	acgttcagat	acccactacc	18360
agtacaccca	gtattgccac	cgccacagag	ggcatggaga	cacaaacgtc	cccggttgcc	18420
tcagcggtgg	cggatgccgc	ggtgcaggcg	gtcgctgcgg	ccgcgtccaa	gacctctacg	18480
gaggtgcaaa	cggacccgtg	gatgtttcgc	gtttcagccc	cccggcgccc	gcgcggttcg	18540
aggaagtacg	gcgcccgcgt	cgcgctactg	cccgaatatg	ccctacatcc	ttccattgcg	18600
cctacccccg	gctatcgctg	ctacacctac	cgccccagaa	gacgagcaac	tacccgacgc	18660
cgaaccacca	ctggaacccg	ccgcgcgcgt	cgccgtcgcc	agcccgtgct	ggccccgatt	18720
tccgtgcgca	gggtggctcg	cgaaggaggc	aggaccctgg	tgctgccaac	agcgcgctac	18780
caccccagca	tcgtttaaaa	gccggtcttt	gtgggtcttg	cagatatggc	cctcacctgc	18840

cgccctccgtt	tcccgggtgcc	gggattccga	ggaagaatgc	accgtaggag	gggcatggcc	18900
ggccacgggcc	tgacggggcgg	catgcgtcgt	gcgaccacc	ggcgggcggcg	cgcgctcgac	18960
cgctcgcatgc	gcggcggtat	cctgcccctc	cttattccac	tgatcgccgc	ggcgattggc	19020
gccgtgcccg	gaattgcate	cgtggccttg	caggcgacga	gacactgatt	aaaaacaagt	19080
tgcatgtgga	aaaatcaaaa	taaaaagctc	ggactctcac	gctcgcttgg	tcctgtaact	19140
atthttgtaga	atggaagaca	tcaactttgc	gtctctggcc	ccgcgacacg	gctcgcgccc	19200
gttcatggga	aactggcaag	atatacggcac	cagcaatatg	agcgggtggcg	ccttcagctg	19260
gggctcgctg	tggagcggca	ttaaaaaattt	cggttccacc	gttaagaact	atggcagcaa	19320
ggcctgggaac	agcagcacag	gccagatgct	gagggataag	ttgaaagagc	aaaatttcca	19380
acaaaagggtg	gtagatggcc	tggcctctgg	cattagcggg	gtgggtggacc	tggccaacca	19440
ggcagtgcaa	aataagatta	acagtaagct	tgatccccgc	cctcccgtag	aggagcctcc	19500
accggcccggtg	gagacagtgt	ctccagaggg	gcgtggcgaa	aagcgtccgc	gccccgacag	19560
ggaagaaact	ctggtgacgc	aaatagacga	gcctccctcg	tacgaggagg	cactaaagca	19620
aggcctgccc	accaccgctc	ccatcgcgcc	catgggtacc	ggagtgtctg	gccagcacac	19680
acccgtaacg	ctggacctgc	ctccccccgc	cgacaccacg	cagaaacctg	tgctgccagg	19740
cccgaaccgcc	gttgtttgtaa	cccgtcctag	ccgcgcgtcc	ctgcgcgcgc	ccgccagcgg	19800
tccgcgatcg	ttgcggccccg	tagccagtgg	caactggcaa	agcacactga	acagcatcgt	19860
gggtctgggg	gtgcaatccc	tgaagcgccg	acgatgcttc	tgaatagcta	acgtgtcgta	19920
tgtgtgtcat	gtatgcgtcc	atgtcgccgc	cagaggagct	gctgagcccg	cgcgcgcccc	19980
ctttccaaga	tggctacccc	ttcgatgatg	ccgcagtggt	cttacatgca	catctcgggc	20040
caggacgcct	cggagtacct	gagccccggg	ctggtgcagt	ttgcccgcgc	caccgagacg	20100
tacttcagcc	tgaataacaa	gtttagaaac	cccacgggtg	cgctacgca	cgacgtgacc	20160
acagaccggt	cccagcgttt	gacgctgcgg	ttcatccctg	tggaccgtga	ggatactcgc	20220
tactcgtaca	aggcgcggtt	caccctagct	gtgggtgata	accgtgtgct	ggacatggct	20280
tccacgtact	ttgacatccg	cggcgtgctg	gacagggggc	ctacttttaa	gccctactct	20340
ggcactgcct	acaacgcctt	ggctcccaag	gggtgcccac	atccttgcca	atgggatgaa	20400
gctgctactg	ctcttgaat	aaacctagaa	gaagaggacg	atgacaacga	agacgaagta	20460
gacgagcaag	ctgagcagca	aaaaactcac	gtatttgggc	aggcgcccta	ttctggtata	20520
aatattacaa	aggagggtat	tcaaataaggt	gtcgaaggct	aaacaccta	atatgccgat	20580
aaaacatttc	aacctgaacc	tcaaataagga	gaatctcagt	ggtacgaaac	tgaattaat	20640
catgcagctg	ggagagtcc	taaaaagact	accccaatga	aaccatgtta	cggttcatat	20700
gcaaaaccca	caaatgaaaa	tggagggcaa	ggcattcttg	taaagcaaca	aatgggaaag	20760
ctagaaagtc	aagtggaaat	gcaatttttc	tcaactactg	aggcgaccgc	aggcaatggt	20820
gataacttga	ctcctaaagt	ggtattgtac	agtgaagatg	tagatataga	aaccccagac	20880
actcatattt	cttacatgcc	cactattaag	gaaggtaact	cacgagaact	aatgggcca	20940
caatctatgc	ccaacaggcc	taattacatt	gcttttaggg	acaattttat	tggctctaat	21000
tattacaaca	gcacgggtaa	tatgggtgtt	ctggcgggcc	aagcatcgca	gttgaatgct	21060
gttgtagatt	tgcaagacag	aaacacagag	ctttcatacc	agcttttgct	tgattccatt	21120
ggtgatagaa	ccagggtactt	ttctatgtgg	aatcaggctg	ttgacagcta	tgatccagat	21180
gttagaatta	ttgaaaatca	tggaaactgaa	gatgaacttc	caaattactg	ctttccactg	21240
ggaggtgtga	ttaatacaga	gactcttacc	aaggtaaaac	ctaaaacagg	tcaggaaaaat	21300
ggatgggaaa	aagatgctac	agaattttca	gataaaaaatg	aaataagagt	tggaaataat	21360
tttgccatgg	aatcaatct	aaatgccaac	ctgtggagaa	atttcctgta	ctccaacata	21420
gcgctgtatt	tgcccagacaa	gctaaagtac	agtccttcca	acgtaaaaat	ttctgataac	21480
ccaaacacct	acgactacat	gaacaagcga	gtgggtggctc	ccgggttagt	ggactgctac	21540
attaaccttg	gagcacgctg	gtcccttgac	tatatggaca	acgtcaacc	atttaaccac	21600
caccgcaatg	ctggcctgcg	ctaccgctca	atgttgctgg	gcaatggctg	ctatgtgccc	21660
ttccacatcc	agggtgcctca	gaagtctctt	gccattaaaa	acctccttct	cctgccgggc	21720
tcatacacct	acgagtggaa	cttcaggaag	gatgttaaca	tggttctgca	gagctcccta	21780
ggaaatgacc	taagggttga	cggagccagc	attaagtttg	atagcatttg	cctttacgcc	21840
accttcttcc	ccatggccca	caacaccgcg	tccagccttg	aggccatgct	tagaaacgac	21900
accaacgacc	agtcctttaa	cgactatctc	tccgccgcca	acatgctcta	ccctataccc	21960
gccaacgcta	ccaacgtgcc	catatccatc	ccctcccgc	actgggcggc	tttccgcggc	22020
tgggccttca	cgcgccttaa	gactaaggaa	accccatcac	tgggctcggg	ctacgaccct	22080
tattacacct	actctggctc	tataccctac	ctagatggaa	ccttttacct	caaccacacc	22140
tttaagaagg	tggccattac	ctttgactct	tctgtcagct	ggcctggcaa	tgaccgcctg	22200
cttaccceca	acgagtttga	aattaagcgc	tcagttgacg	gggaggggta	caacggtgcc	22260
cagtgttaaca	tgaccaaaga	ctgggttcctg	gtacaaatgc	tagctaacta	caacattggc	22320
taccagggct	tctatatccc	agagagctac	aaggaccgca	tgtactcctt	cttttagaaac	22380
ttccagccca	tgagccgtca	ggtgggtgac	gatactaaat	acaaggacta	ccaacaggtg	22440
ggcatcctac	accaacacaa	caactctgga	tttgttggct	accttgcccc	caccatgcgc	22500

-21-

gaaggacagg	cctaccctgc	taacttcccc	tatccgctta	taggcaagac	cgcagttgac	22560
agcattaccc	agaaaaagtt	tctttgcat	cgcacccttt	ggcgcacccc	attctccagt	22620
aactttatgt	ccatgggccc	actcacagac	ctggggccaaa	accttctcta	cgccaactcc	22680
gcccacgcgc	tagacatgac	ttttgaggtg	gatcccatgg	acgagcccac	ccttctttat	22740
gttttgtttg	aagtctttga	cgtgggccgt	gtgcaccggc	cgcaccgcgg	cgtcatcgaa	22800
accgtgtacc	tgcgcacgcc	cttctcggcc	ggcaacgcca	caacataaag	aagcaagcaa	22860
catcaacaac	agctgccgcc	atgggctcca	gtgagcagga	actgaaagcc	attgtcaaag	22920
atcttggttg	tggggccatat	tttttgggca	cctatgacaa	gcgctttcca	ggctttgttt	22980
ctccacacaa	gctcgcctgc	gccatagtca	atacggccgg	tcgcgagact	gggggcgtac	23040
actggatggc	ctttgcctgg	aaccgcact	caaaaacatg	ctacctcttt	gagccctttg	23100
gcttttctga	ccagcgactc	aagcaggttt	accagtttga	gtacgagtca	ctcctgcgcc	23160
gtagcgccat	tgtctcttcc	cccgaccgct	gtataacgct	ggaaaagtcc	acccaaagcg	23220
tacagggggc	caactcggcc	gcctgtggac	tattctgctg	catgtttctc	cacgcctttg	23280
ccaactggcc	ccaaactccc	atggatcaca	accccaccat	gaaccttatt	accggggtag	23340
ccaactccat	gctcaacagt	ccccaggtag	agcccaccct	gcgtcgcaac	caggaacagc	23400
tctacagctt	cctggagcgc	cactcgccct	acttccgcag	ccacagtgcg	cagattagga	23460
gcgccacttc	tttttgtcac	ttgaaaaaca	tgtaaaaata	atgtactaga	gacactttca	23520
ataaaggcaa	atgcttttat	ttgtacactc	tcgggtgatt	atttaccctc	acccttgccg	23580
tctgcgccgt	ttaaaaatca	aaggggttct	gccgcgcac	gctatgcgcc	actggcaggg	23640
acacgttgcg	atactggtgt	ttagtgctcc	acttaaactc	aggcacaacc	atccgcggca	23700
gctcggtgaa	gttttctact	cacaggtgc	gcaccatcac	caacgcgttt	agcaggtcgg	23760
gcgccgatata	cttgaaagtgc	cagttggggc	ctccgccttg	cgcgcgcgag	ttgcgatata	23820
cagggttgca	gcactggaac	actatcagcg	ccgggtggtg	cacgctggcc	agcacgctct	23880
tgtcggagat	cagatcccg	tcaggtcct	ccgcgttgct	cagggcgaa	ggagtcaact	23940
ttggtagctg	ccttcccaaa	aagggcgcg	gcccaggctt	tgagttgcac	tcgcaccgta	24000
gtggcatcaa	aaggtgaccg	tgcccgtct	gggcgttagg	atacagcgcc	tgcatataaag	24060
ccttgatctg	cttaaaagcc	acctgagcct	ttgcgccttc	agagaagaac	atgccgcaag	24120
acttgccgga	aaactgattg	gccggacagg	ccgcgtcgctg	cacgcagcac	cttgcgtcgg	24180
tgttgagat	ctgcaccaca	tttcggcccc	accggttctt	cacgatcttg	gccttgctag	24240
actgctcctt	cagcgcgcg	tgcccgtttt	cgctcgtcac	atccatttca	atcacgtgct	24300
ccttattttat	cataatgctt	ccgtgtagac	acttaagctc	gccttcgatc	tcagcgcgag	24360
ggtgcagcca	caacgcgcag	cccgtgggct	cgtgatgctt	gtaggtcacc	tctgcaaacg	24420
actgcagga	cgctcgcagg	aatcgcccc	tcactcgtcac	aaaggtcttg	ttgctgggtga	24480
aggtcagctg	caaccgcgg	tgctcctcgt	tcagccagggt	cttgcatagc	gccgccagag	24540
cttccacttg	gtcaggcagt	agtttgaagt	tcgccttttag	atcgttatcc	acgtggtaact	24600
tgtccatcag	cgcgcgcgca	gcctccatgc	ccttctccca	cgagacacg	atcggcacac	24660
tcagcggggt	catcacccga	atttcaactt	ccgctctcgt	gggctcttcc	tcttctctct	24720
gcctcgcgat	accacgcgcc	actgggtcgt	cttcattcag	ccgcgcgact	gtgcgcttac	24780
ctcctttgcc	atgcttgatt	agcaccggtg	ggttgctgaa	accaccatt	tgtagcgcca	24840
catcttctct	ttcttctctg	ctgtccacga	ttacctctgg	tgatggcggg	cgctcgggct	24900
tgggagaagg	gcgcttcttt	ttcttcttgg	gcgcaatggc	caaatccgcc	gccgaggtcg	24960
atggccgcgg	gctgggtgtg	cgcggcacca	gcgcgtcttg	tgatgagtct	tcctcgtcct	25020
cggactcgat	acgcgcctc	atccgctttt	ttgggggcgc	ccggggaggc	ggcggcgacg	25080
gggacgggga	cgacacgtcc	tccatgggtg	ggggacgtcg	cgccgcaccg	cgtccgcgct	25140
cgggggtggt	ttcgcgctgc	tctcttcccc	gactggccat	ttccttctcc	tataggcaga	25200
aaaagatcat	ggagtcagtc	gagaagaagg	acagcctaac	cgccccctct	gagttcgcca	25260
ccaccgcctc	caccgatgcc	gccaacgcgc	ctaccacctt	ccccgtcgag	gcacccccgc	25320
ttgaggagga	ggaagtgatt	atcgagcagg	acccaggttt	tgtaagcgaa	gacgacgagg	25380
accgctcagt	accaacagag	gataaaaagc	aagaccagga	caacgcagag	gcaaacgagg	25440
aacaagtccg	gcgggggggac	gaaaggcatg	gcgactacct	agatgtggga	gacgacgtgc	25500
tgttgaagca	tctgcagcgc	cagtgcgcc	ttatctgcga	cgcggttgaa	gagcgacgag	25560
atgtgccct	cgccatagcg	gatgtcagc	ttgcctacga	acgccacctc	ttctcaccgc	25620
gcgtaccccc	caaacgcgca	gaaaacggca	catgcgagcg	caaccgcgc	ctcaacttct	25680
accccgattt	tgcogtgcca	gaggtgcttg	ccacctatca	catctttttc	caaaactgca	25740
agataccct	atcctgccgt	gccaaccgca	gccgagcgga	caagcagctg	gccttgcggc	25800
agggcgctgt	catacctgat	atcgctcgc	tcaacgaagt	gccaaaaatc	tttgagggtc	25860
ttggacgcga	cgagaagcgc	gcggcaaacg	ctctgcaaca	ggaaaaacagc	gaaaatgaaa	25920
gtcactctgg	agtgttggtg	gaactcgagg	gtgacaacgc	gcgcctagcc	gtactaaaac	25980
gcagcatcga	ggtcaccac	tttgccctacc	cggcacttaa	cctaccccc	aaggtcatga	26040
gcacagtcac	gagtgagctg	atcgctgcgc	gtgcgcagcc	cctggagagg	gatgcaaatt	26100
tgcaagaaca	aacagaggag	ggcctaccgc	cagttggcga	cgagcagcta	gcgcgctggc	26160

-22-

ttcaaaccg	cgagcctgcc	gacttggagg	agcgacgcaa	actaatgatg	gccgcagtgc	26220
tcgttaccgt	ggagcttgag	tgcattgcagc	gggtcttttgc	tgaccggag	atgcagcgca	26280
agctagagga	aacattgcac	tacaccttcc	gacagggcta	cgtacgccag	gcctgcaaga	26340
tctccaacgt	ggagctctgc	aacctgggtct	cctaccttgg	aatttttgcac	gaaaaccgcc	26400
ttgggcaaaa	cgtgcttcat	tccacgctca	agggcgaggc	gcgcccgcac	tacgtccgcg	26460
actgcttcta	cttatttcta	tgctacacct	ggcagacggc	catgggcgtt	tgccagcagt	26520
gcttggagga	gtgcaacctc	aaggagctgc	agaaactgct	aaagcaaaaac	ttgaaggacc	26580
tatggacggc	cttcaacgag	cgctccgtgg	ccgcgcacct	ggcgacacac	atcttccccc	26640
aacgcctgct	taaaaccctg	caacagggtc	tgccagactt	caccagtcaa	agcatgttgc	26700
agaacttttag	gaactttatc	ctagagcgct	caggaatctt	gcccggccacc	tgctgtgcac	26760
ttcctagcga	ctttgtgccc	attaagtacc	gcgaatgccc	tccgcccgtt	tgggggccact	26820
gctaccttct	gcagctagcc	aactaccttg	cctaccactc	tgacataatg	gaagacgtga	26880
gcggtgacgg	tctactggag	tgctactgtc	gctgcaacct	atgcaccccc	caccgctccc	26940
tggtttgcaa	ttcgcagctg	cttaacgaaa	gtcaaattat	cggtaccttt	gagctgcagg	27000
gtccctcgcc	tgacgaaaag	tccgcggtct	cggggttgaa	actcactccg	gggctgtgga	27060
cgctcggtta	ccttcgcaaa	tttgtacctg	aggactacca	cgcccacgag	attaggttct	27120
acgaagacca	atcccggccc	ccaaatgcgg	agcttaccgc	ctgogtcatc	acccagggcc	27180
acattcttgg	ccaattgcaa	gccaatcaaca	aagcccgcga	agagtctctg	ctacgaaagg	27240
gacggggggg	ttacttggac	ccccagtcgg	gcgaggagct	caacccaatc	cccccgccgc	27300
cgagcccta	tcagcagcag	ccgcccggccc	ttgcttccca	ggatggcacc	caaaaagaag	27360
ctgcagctgc	cgccggccacc	cacggacgag	gaggaatact	gggacagtca	ggcagaggag	27420
gttttggacg	aggaggagga	ggacatgatg	gaagactggg	agagcctaga	cgaggaagct	27480
tccgaggtcg	aagaggtgtc	agacgaaaca	ccgtcacctt	cggtcgcatc	cccctcgccg	27540
gcgcccacga	aatcggcaac	cggttccagc	atggctacaa	cctccgctcc	tcaggcgccg	27600
ccggcactgc	ccgttcgccc	acccaaccgt	agatgggaca	ccactggaac	cagggccggt	27660
aagtccaagc	agccgcccgc	gttagcccaa	gagcaacaac	agcgccaagg	ctaccgctca	27720
tgccgcccgc	acaagaacgc	catagtgtct	tgcttgcaag	actgtggggg	caacatctcc	27780
ttcggcccgc	gctttcttct	ctaccatcac	ggcgtggcct	tccccgtaaa	catcctgcat	27840
tactaccgtc	atctctacag	cccatactgc	accggcggca	gcggcagcgg	cagcaacagc	27900
agcggccaca	cagaagcaaa	ggcgaccgga	tagcaagact	ctgacaaaag	ccaagaaatc	27960
cacagcgccg	gcagcagcag	gaggaggagc	gctgcgtctg	gcgcccacac	aacccgtatc	28020
gacccgagag	cttagaaaaca	ggatttttcc	cactctgtat	gctatatatt	aacagagcag	28080
gggccaagaa	caagagctga	aaataaaaaa	caggtctctg	cgatccctca	cccgcagctg	28140
cctgtatcac	aaaagcgaag	atcagcttcc	gcgcacgctg	gaagacgcgg	aggctctctt	28200
cagtaaatca	tgccgcttga	ctcttaagca	ctagtcttgc	gccctttctc	aaatttaagc	28260
gcgaaaacta	cgtcatctcc	agcggccaga	ccggcgccca	gcacctgtcg	tcagcgccat	28320
tatgagcaag	gaaattccca	cgccctacat	gtggagttac	cagccacaaa	tgggacttgc	28380
ggctggagct	gccaagactc	actcaacccg	aataaactac	atgagcgccg	gacccacatc	28440
gatatcccg	gtcaacggaa	tccgcgcccc	ccgaaaaccga	attctcttgg	aacaggcgcc	28500
tattaccacc	acacctcgta	ataaccttaa	tccccgtagt	tgcccccgtg	ccctggtgta	28560
ccaggaaagt	cccgctccca	ccactgttgt	acttcccaga	gacgcccagg	ccgaagtcca	28620
gatgactaac	tcagggggcg	agcttgcggg	cggtttctgt	cacagggtgc	ggctcgcccg	28680
gcaggggtata	actcacctga	caatcagagg	gcgaggtatt	cagctcaacg	acgagtcggt	28740
gagctcctcg	cttggctctc	gtccggacgg	gacatttcag	atcgccggcg	ccggccgctc	28800
ttcattcacg	cctcgtcagg	caatcctaac	tctgcagacc	tcgtcctctg	agccgcgctc	28860
tgagggcatt	ggaactctgc	aattttattga	ggagtttgtg	ccatcggtct	actttaaccc	28920
cttctcgggg	cctcccggcc	actatccgga	tcaatttatt	cctaactttg	acgcggtaaa	28980
ggactcggcg	gacggctacg	actgaatgtt	aagtggagag	gcagagcaac	tgccgctgaa	29040
acacctgggt	cactgtcgcc	gccacaagtg	ctttgcccgc	gactccgggt	agtttttgta	29100
ctttgaattg	cccgaggatc	atatcgaggg	cccgccgcac	ggcgtccggc	ttaccgcca	29160
gggagagctt	gcccgtagcc	tgattcggga	gtttaccacc	cgccccctgc	tagttgagcg	29220
ggacagggga	ccctgtgttc	tcactgtgat	ttgcaactgt	cctaactttg	gattacatca	29280
agatctttgt	tgccatctct	gtgctgagta	taataaatac	agaaattaaa	atatactggg	29340
gctcctatcg	ccatcttgta	aacggccacc	tcttcacccc	ccaagcaaaa	ccaaggcgaa	29400
ccttacctgg	tacttttaac	atctctccct	ctgtgattta	caacagtttc	aacccagacg	29460
gagtgagtct	acgagagaa	ctctccgagc	tcagctactc	catcagaaaa	aacaccaccc	29520
tccttacctg	ccgggaacgt	acgagtgctg	caccggccgc	tgaccacac	ctaccgctg	29580
accgtaaac	agactttttc	cggacagacc	tcaataactc	tgtttaccag	aacaggaggt	29640
gagcttagaa	aacctttagg	gtattaggcc	aaaggcgacg	ctactgtggg	gtttatgaac	29700
aattcaagca	actctacggg	ctatttcaat	tcaggtttct	ctagaaatgg	acggaattat	29760
tacagagcag	cgcctgctag	aaagacgcag	ggcagcggcc	gagcaacagc	gcatgaatca	29820

agagctccaa	gacatgggta	acttgaccca	gtgcaaaaagg	ggtatctttt	gtctgggtaaa	29880
gcaggccaaa	gtcacctacg	acagtaatac	caccggacac	cgccttagct	acaagttgcc	29940
aaccaagcgt	cagaaattgg	tggtcatggg	gggagaaaag	cccattacca	taactcagca	30000
ctcggtagaa	accgaaggct	gcattcaactc	accttgtaaa	ggacctgagg	atctctgcac	30060
ccttattaag	accctgtgcg	gtctcaaaga	tcttattccc	tttaactaat	aaaaaaaaat	30120
aataaagcat	cacttactta	aaatcagtta	gcaaatttct	gtccagttta	ttcagcagca	30180
cctccttgcc	ctcctcccag	ctctgggtatt	gcagcttcct	cctggctgca	aactttctcc	30240
acaatctaaa	tgggaatgtca	gtttcctcct	gttcctgtcc	atccgcaccc	actatcttca	30300
tgttggtgca	gatgaagcgc	gcaagaccgt	ctgaagatac	cttcaacccc	gtgtatccat	30360
atgacacgga	aaccggtcct	ccaactgtgc	ctttcttac	tcctcccttt	gtatccccc	30420
atggggttca	agagagtccc	cctgggggtac	tctctttgcg	cctatccgaa	cctctagtta	30480
cctccaatgg	catgcttgcg	ctcaaaatgg	gcaacggcct	ctctctggac	gaggccggca	30540
accttacctc	ccaaaatgta	accactgtga	gcccacctct	caaaaaaacc	aagtcaaaca	30600
taaacctgga	aatatctgca	ccccctacag	ttacctcaga	agccctaact	gtggctgccg	30660
ccgcacctct	aatgggtcgcg	ggcaacacac	tcacctgca	atcacaggcc	ccgctaaccg	30720
tgcacgactc	caaacttagc	attgccaccc	aaggaccct	cacagtgtca	gaaggaaagc	30780
tagccctgca	aacatcaggc	cccctcacca	ccaccgatag	cagtaccctt	actatcactg	30840
cctcaccccc	tctaactact	gccactggta	gcttgggcat	tgacttgaaa	gagcccat	30900
atacacaaaa	tggaaaacta	ggactaaaag	acggggctcc	tttgcatgta	acagacgacc	30960
taaacacttt	gaccgtagca	actgggtccag	gtgtgactat	taataaact	tccttgcaaa	31020
ctaaagttac	tggagccttg	ggttttgatt	cacaaggcaa	tatgcaactt	aatgtagcag	31080
gaggactaag	gattgattct	caaaacagac	gccttatact	tgatgttagt	tatccgtttg	31140
atgctcaaaa	ccaactaaat	ctaagactag	gacagggccc	tctttttata	aactcagccc	31200
acaacttggg	tattaactac	aacaaaggcc	ttactttgtt	tacagcttca	aacaattcca	31260
aaaagcttga	ggttaacctta	agcactggca	aggggttgat	gtttgacgct	acagccatag	31320
ccattaatgc	aggagatggg	cttgaatttg	gttcacctaa	tgcaccaaac	acaaatcccc	31380
tcaaaacaaa	aattggccat	ggcctagaat	ttgattcaaa	caaggctatg	gttcctaaac	31440
taggaactgg	ccttagtttt	gacagcacag	gtgccattac	agtaggaaac	aaaaataatg	31500
ataagctaac	tttgtggacc	acaccagctc	catctcctaa	ctgtagacta	aatgcagaga	31560
aagatgctaa	actcactttg	gtcttaacaa	aatgtggcag	tcaaataactt	gctacagttt	31620
cagttttggc	tgttaaaggc	agtttggctc	caatatctgg	aacagttcaa	agtgtcatc	31680
ttattataag	atgtgacgaa	aatggagtgc	tactaaacaa	ttccttctctg	gaccagaat	31740
attggaactt	tagaaatgga	gatcttactg	aaggcacagc	ctatacaaac	gctgttggat	31800
ttatgcctaa	cctatcagct	tatccaaaat	ctcacggtaa	aactgccaaa	agtaacattg	31860
tcagtcaagt	ttacttaaac	ggagacaaaa	ctaaacctgt	aacactaacc	attacactaa	31920
acgggtacaca	ggaaacagga	gacacaactc	caagtgcata	ctctatgtca	ttttcatggg	31980
actgggtctgg	ccacaactac	attaatgaaa	tatttgccac	atcctcttac	actttttcat	32040
acattgcccc	agaataaaga	atcgtttgtg	ttatgtttca	acgtgtttat	ttttcaattg	32100
cagaaaattt	caagtcattt	ttcattcagt	agtatacccc	caccaccaca	tagcttatac	32160
agatcacctg	accttaatca	aactcacaga	accctagtat	tcaacctgcc	acctccctcc	32220
caacacacag	agtacacagt	cctttctccc	cggctggcct	taaaaagcat	catatcatgg	32280
gtaacagaca	tattcttagg	tgttatattc	cacacggttt	cctgtcgagc	caaacgctca	32340
tcagtgatat	taataaactc	cccgggcagc	tcacttaagt	tcatgtcgct	gtccagctgc	32400
tgagccacag	gctgctgtcc	aacttgcggt	tgcttaacgg	gcggcgaagg	agaagtccac	32460
gocctacatgg	gggtagagtc	ataatcgtgc	atcaggatag	ggcgggtggtg	ctgcagcagc	32520
gcgcgaataa	actgctgcgc	ccgccgctcc	gtcctgcagg	aatacaacat	ggcagtggtc	32580
tcctcagcga	tgattcgcac	cgccccgagc	ataaggcgcc	ttgtcctccg	ggcacagcag	32640
cgcacctgga	tctcacttaa	atcagcacag	taactgcagc	acagcaccac	aataattgtc	32700
aaaatcccac	agtgaaggc	gctgtatcca	aagctcatgg	cggggaccac	agaaccacg	32760
tggccatcat	accacaagcg	caggtagatt	aagtggcgac	ccctcataaa	cacgctggac	32820
ataaacatta	cctcttttgg	catgttgtaa	ttcaccacct	cccggtacca	tataaacctc	32880
tgattaaaca	tggcgccatc	caccaccatc	ctaaaccagc	tggccaaaaac	ctgcccgcgcg	32940
gctatacact	gcagggaacc	gggactggaa	caatgcagtc	ggagagccca	ggactcgtaa	33000
ccatggatca	tcattgctcgt	catgatatca	atgttggtcac	aacacaggca	cacgtgcata	33060
cacttctctca	ggattacaag	ctcctccgcg	gttagaacca	tatcccaggg	aacaacccat	33120
tctgaatca	gcgtaaatcc	cacactgcag	ggaagacctc	gcacgtaact	cacgttggtgc	33180
attgtcaaaag	tgttacattc	gggcagcagc	ggatgacctc	ccagtatggt	agcgcggtt	33240
tctgtctcaa	aaggaggtag	acgatcccta	ctgtacggag	tcgcccagaga	caaccgagat	33300
cgtgttggtc	gtagtgtcat	gccaaatgga	acgcccagcg	tagtcatatt	tcctgaagca	33360
aaaccaggtg	cgggcgtgac	aaacagatct	gcgtctccgg	tctcgccgct	tagatcgctc	33420
tgtgtagtag	ttgtagtata	tccactctct	caaagcatcc	aggcgccccc	tggcttcggg	33480

-24-

```

ttctatgtaa actccttcat ggcgcgctgc cctgataaca tccaccaccg cagaataagc 33540
cacaccacagc caacctacac attcgttctg cgagtcacac acgggaggag cggaagagc 33600
tggaagaacc atgttttttt ttttattcca aaagattatc caaaacctca aaatgaagat 33660
ctattaagtg aacgcgctcc cctccgggtg cgtggtcaaa ctctacagcc aaagaacaga 33720
taatggcatt tgtaagatgt tgcacaatgg cttccaaaag gcaaaccggcc ctcacgtcca 33780
agtggacgta aaggctaaac ccttcagggt gaatctcctc tataaacatt ccagcacctt 33840
caaccatgcc caaataattc tcatctcgcc accttctcaa tatatctcta agcaaatccc 33900
gaatattaag tccggccatt gtaaaaatct gctccagagc gccctccacc ttcagcctca 33960
agcagcgaat catgattgca aaaattcagg ttcttcacag acctgtataa gattcaaaag 34020
cggaacatta acaaaaatac cgcgatcccg taggtccctt cgcagggccca gctgaacata 34080
atcgtgcagg tctgcacgga ccagcgcggc cacttccccg ccaggaaacct tgacaaaaga 34140
acccacactg attatgacac gcatactcgg agctatgcta accagcgtag ccccgatgta 34200
agctttgttg catggggcggc gatataaaat gcaagggtgt gctcaaaaaa tcaggcaaag 34260
cctcgcgcaa aaaagaaaagc acatcgtagt catgctcatg cagataaaagg caggtaagct 34320
ccggaaccac cacagaaaaa gacaccattt ttctctcaaa catgtctgcg ggtttctgca 34380
taaacacaaa ataaaataac aaaaaaacat ttaaacatta gaagcctgtc ttacaacagg 34440
aaaaacaacc cttataagca taagacggac tacggccatg ccggcgtgac cgtaaaaaaa 34500
ctggtcaccc tgattaaaaa gcaccaccga cagctcctcg gtcattgtccg gagtcataat 34560
gtaagactcg gtaaacacat caggttgatt catcggtcag tgctaaaaag cgaccgaaat 34620
agcccggggg aatacatacc cgcaggcgta gagacaacat tacagccccc ataggaggta 34680
taacaaaatt aataggagag aaaaacacat aaacacctga aaaacctcc tgcctaggca 34740
aaatagcacc ctcccgctcc agaacaacat acagcgcttc cacagcggca gccataacag 34800
tcagccttac cagtaaaaaa gaaaacctat taaaaaaaca cactcgaca cggcaccagc 34860
tcaatcagtc acagtgtaaa aaagggccaa gtgcagagcg agtatatata ggactaaaaa 34920
atgacgtaac ggttaaagtc cacaaaaaac acccagaaaa ccgcacgcga acctacgccc 34980
agaaacgaaa gccaaaaaac ccacaacttc ctcaaactgt cacttcggtt tccccacgtt 35040
acgtcacttc ccattttaat taagaaaact acaattccca acacatacaa gttactccgc 35100
cctaaaacct acgtcacccg ccccggtccc acgccccgcg ccacgtcaca aactccacce 35160
cctcattatc atattggctt caatccaaaa taaggatatat tattgatgat g 35211

```

<210> 44

<211> 33622

<212> DNA

<213> Artificial Sequence

<220>

<223> Plasmid Av3nBg

<400> 44

```

catcatcaat aatatacctt attttggatt gaagccaata tgataatgag ggggtggagt 60
ttgtgacgtg ggcggyggcg tgggaacggg gcggtgacg tagtagtgtg gcggaagtgt 120
gatgttgcaa gtgtggcgga acacatgtaa gcgacggatg tggcaaaagt gacgtttttg 180
gtgtgcgccg gtgtacacag gaagtgacaa ttttcgcgcg gttttaggcg gatgtttagt 240
taaatttggg cgtaaccgag taagatttgg ccattttcgc gggaaaactg aataagagga 300
agtgaatctt gaataatttt gtgttactca tagcgcgtaa tatttgtcta gggccgcggg 360
gactttgacc gtttacgtgg agactcgccc agggcgcgcc ccgatgtacg ggccagatat 420
acgcgtatct gaggggacta ggggtgtggtt agggcgaagg cggggcttcg gttgtacgcy 480
gttaggagtc cctcaggat atagtagttt cgcttttgca tagggagggg gaaatgtagt 540
cttatgcaat actctttagt tcttgcaaca tggtaacgat gagttagcaa catgccttac 600
aaggagagaa aaagcaccgt gcatgccgat tgggtggaag aaggtggtac gatcgtgcct 660
tattaggaag gcaacagacg ggtctgacat ggattggacg aaccactgaa ttccgcattt 720
cagagatatt gtattttaagt gcctagctcg atacaataaa cgccatttga ccattcacca 780
cattgggtgtg cacctccggc cctggccact ctcttcgcga tcgctgtctg cgggggcccag 840
ctgttgggct cgcggttgag gacaaactct tcgcgggtctt tccagtaact ttggatcgga 900
aaccgctcgg cctccgaacg gtaactccgc gccaggggac ctgagcgagt ccgcatcgac 960
cggatcgga aacctctcga gaaaggcgtg taaccagtca cagtcgctct agaactagt 1020
gatcccccg gctgcaggaa ttcgatctag atggataaag gtccaaaaaa gaagagaaag 1080
gtagaagacc ccaaggactt tccttcagaa ttgctaagtt ttttgagtga ttactggcc 1140
gtcgttttcc aacgtcgtga ctgggaaaaa cctggcggtta cccaacttaa tcgccttgca 1200
gcacatcccc ctttcgccag ctggcgtaat agcgaagagg cccgcaccga tcgcccttcc 1260
caacagttgc gcagcctgaa tggcgaatgg cgctttgcct ggtttccggc accagaagcg 1320

```


-25-

gtgcccggaaa	gctgggctgga	gtgcgatctt	cctgaggccg	atactgtcgt	cgccccctca	1380
aactggcaga	tgcacgggta	cgatgcgccc	atctacacca	acgtaacct	tcccatctacg	1440
gtcaatccgc	cgttttgtcc	cacggagaat	ccgacggggt	gttactcgct	cacatttaaat	1500
gttgatgaaa	gctggctaca	ggaaggccag	acgcgaatta	tttttgatgg	cgtaaactcg	1560
gcgtttcatc	tgtggtgcaa	cgggcgctgg	gtcgggttacg	gccaggacag	tcgtttgccg	1620
tctgaatttg	acctgagcgc	attttttacgc	gccggagaaa	accgcctcgc	ggtgatgggtg	1680
ctgcgttgga	gtgacggcag	ttatctggaa	gatcaggata	tgtggcggat	gagcggcatt	1740
ttccgtgacg	tctcggtgct	gcataaacccg	actacacaaa	tcagcgattt	ccatggtgcc	1800
actcgcttta	atgatgattt	cagccgcgct	gtactggagg	ctgaagttca	gatgtgcggc	1860
gagttgcgtg	actacctacg	ggtaacagtt	tctttatggc	agggtgaaac	gcaggtcgcc	1920
agcggcaccg	cgcctttcgg	cggtgaaatt	atcgatgagc	gtggtgggtta	tgccgatcgc	1980
gtcacactac	gtctgaacgt	cgaaaacccg	aaactgtgga	gcgcggaaat	cccgaatctc	2040
tatcgtagcg	tggttgaaact	gcacaccgcc	gacggcacgc	tgattgaagc	agaagcctgc	2100
gatgtcgggt	tccgcgaggt	gcggattgaa	aatggtctgc	tgctgctgaa	cggaagcccg	2160
ttgctgattc	gaggcgtaa	ccgtcacag	catcatctc	tgcatggtca	ggtcatggat	2220
gagcagacga	tggtgcagga	tatcctgctg	atgaagcaga	acaacttta	cgccgtgcgc	2280
tgttcgcatt	atccgaacca	tccgctgtgg	tacacgctgt	gcgaccgcta	cgccgtgtat	2340
gtggtggatg	aagccaatat	tgaaacccac	ggcatggtgc	caatgaatcg	tctgaccgat	2400
gatccgcgct	ggctaccggc	gatgagcgaa	cgcgtaacgc	gaatgggtgca	gcgcgatcgt	2460
aatcacccga	gtgtgatcat	ctggtcgctg	gggaatgaat	caggccacgg	cgtaatcac	2520
gacgcgctgt	atcgctggat	caaactgttc	gatccttccc	gcccgggtgca	gtatgaaggc	2580
ggcggagccg	acaccacggc	caccgatatt	atttgcccga	tgtacgcgcg	cgtaggatgaa	2640
gaccagccct	tcccggctgt	gccgaaatgg	tccatcaaaa	aatggctttc	gctacctgga	2700
gagacgcgcc	cgctgatcct	ttgcgaatac	gcccacgcga	tggttaacag	tcttgccggt	2760
ttcgctaaat	actggcaggc	gtttcgctag	tatccccgtt	tacagggcgg	cttcgtctgg	2820
gactgggtgg	atcagtcgct	gattaaatat	gatgaaaacg	gcaaccctgtg	gtcggcttac	2880
ggcgggtgatt	ttggcgatac	gccgaacgat	cgccagttct	gtatgaacgg	tctggtcttt	2940
gccgaccgca	cgccgcatac	agcgcgtgacg	gaagcaaaac	accagcagca	gtttttccag	3000
ttccgtttat	ccgggcaaac	catcgaagtg	accagcgaat	acctgttccg	tcatagcgat	3060
aacgagctcc	tgacttgat	ggtggcgctg	gatggttaagc	cgctggcaag	cggtgaagtg	3120
cctctggatg	tcgctccaca	aggtaaacag	ttgattgaac	tgccctgaact	accgcagccg	3180
gagagcgccg	ggcaactctg	gctcacagta	cgcgtagtgc	aaccgaacgc	gaccgcattg	3240
tcagaagccg	ggcacatcag	cgcctggcag	cagtggcgctc	tgccggaaaaa	cctcagtgtg	3300
acgctccccg	ccgcgtccca	cgccatcccc	catctgacca	ccagcgaaat	ggatttttgc	3360
atcgagctgg	tgataaagcg	ttggcaattt	aaccgccagt	caggctttct	ttcacagatg	3420
tggtatggcg	ataaaaaaca	actgctgacg	ccgctgcgcg	atcagttcac	ccgtgcaccg	3480
ctggataaacg	acattggcgt	aagtgaagcg	acccgcattg	accctaacgc	ctgggtcgaa	3540
cgctggaagg	cgccggggcca	ttaccaggcc	gaagcagcgt	tggtgcagtg	cacggcagat	3600
acacttgcgt	atgcggtgct	gattacgacc	gctcacgcgt	ggcagcatca	ggggaaaaac	3660
ttattttatca	ccgggaaaac	ctaccggatt	gatggttagtg	gtcaaatggc	gattaccgtt	3720
gatgttgaaag	tgccgagcga	tacaccgcac	ccggcgcgga	ttggcctgaa	ctgccagctg	3780
gcgcaggtag	cagagcgggt	aaactggctc	ggattagggc	cgcaagaaaa	ctatcccgcac	3840
cgccttactg	ccgcctgttt	tgaccgctgg	gatctgccat	tgctcagacat	gtataccccg	3900
tacgtcttcc	cgagcgaaaa	cggtctgcgc	tgccggagcgc	gcgaattgaa	ttatggccca	3960
caccagtggc	gcggcgactt	ccagttcaac	atcagccgct	acagtcaaca	gcaactgatg	4020
gaaaccagcc	atcgccatct	gctgcacgcg	gaagaaggca	catggctgaa	tatcgacggt	4080
ttccatatgg	ggattgggtg	cgacgactcc	tggagcccg	cagtatccgg	ggaatttcag	4140
ctgagcgtcc	gtcgctacca	ttaccagttg	gtctggtgtc	aaaaataata	atctcgaatc	4200
aagcttatcg	ataccgtcga	aacttgttta	ttgcagctta	taatggttac	aaataaagca	4260
atagcatcac	aaatttcaca	aataaagcat	ttttttcact	gcattctagt	tgtggtttgt	4320
ccaaactcat	caatgtatct	tatcatgtct	ggatccgacc	tcggatctgg	aaggtgctga	4380
ggtacgatga	gaccgcgacc	aggtgcagac	cctgcgagtg	tgccgggtaaa	catattagga	4440
accagcctgt	gatgctggat	gtgaccgagg	agctgaggcc	cgatcacttg	gtgctggcct	4500
gcaccgcgcg	tgagtttgcc	tctagcagatg	aagatacaga	ttgaggtact	gaaatgtgtg	4560
ggcgtggctt	aagggtggga	aagaatatat	aagggtgggg	tcttatgtag	ttttgtatct	4620
gttttgcagc	agccgcgcgc	gccatgagca	ccaactcggt	tgatggaagc	attgtgagct	4680
catatttgac	aacgcgcgatg	ccccatggg	ccgggtgtgcg	tcagaatgtg	atgggctcca	4740
gcattgatgg	tcgcccgcgtc	ctgcccgcga	actctactac	cttgacctac	gagaccgtgt	4800
ctggaacgcc	gttgagagact	gcagcctccg	ccgcgccttc	agccgctgca	gccaccgccc	4860
gcgggattgt	gactgacttt	gctttcctga	gcccgccttc	aagcagtgca	gcttcccgtt	4920
catccgcccc	cgatgacaa	ttgacggctc	ttttggcaca	attggattct	ttgaccggg	4980

-26-

aacttaaatgt	cgttttctcag	cagctgttgg	atctgcgccca	gcagggtttct	gccctgaagg	5040
cttcctcccc	tcccaatgcg	gtttaaaaca	taaataaaaa	accagactct	gtttggattt	5100
ggatcaagca	agtgtcttgc	tgtctttatt	taggggtttt	gcgcgcgcgg	tagggccggg	5160
accagcggtc	tcggtcggtg	agggctcctgt	gtattttttc	caggacgtgg	taaagggtgac	5220
tctggatggt	cagatacatg	ggcataagcc	cgtctctggg	gtggaggtag	caccactgca	5280
gagcttcatg	ctgcgggggtg	gtgttgtaga	tgatccagtc	gtagcaggag	cgctgggcgt	5340
gggtgcctaaa	aatgtctttc	agtagcaagc	tgattgccag	gggcaggccc	ttgggtgtaag	5400
tgtttacaaa	gcgggttaagc	tgggatgggt	gcatacgtgg	ggatatgaga	tgcatcttgg	5460
actgtatttt	taggttgggt	atgttcccag	ccatatccct	ccggggattc	atgttgtgca	5520
gaaccaccag	cacagtgtat	ccggtgcact	tgggaaattt	gtcatgtagc	ttagaaggaa	5580
atgctgtggaa	gaacttggag	acgcccttgt	gacctccaag	attttccatg	cattcgtcca	5640
taatgatggc	aatgggcccc	cgggcggcgg	cctgggcgaa	gatattttctg	ggatcactaa	5700
cgtcatagtt	gtgttccagg	atgagatcgt	cataggccat	ttttacaaaag	cgcgggcgga	5760
gggtgccaga	ctgcggtata	atggttccat	ccggcccagg	ggcgtagtta	ccctcacaga	5820
tttgcatttc	ccacgcctttg	agttcagatg	gggggatcat	gtctacctgc	ggggcgatga	5880
agaaaacggt	ttccggggta	ggggagatca	gctgggaaga	aagcaggttc	ctgagcagct	5940
gcgacttacc	gcagccgggtg	ggcccgtaaa	tcacacctat	taccggctgc	aactggtagt	6000
taagagagct	gcagctgcgg	tcacccctga	gcaggggggc	cacttcgtta	agcatgtccc	6060
tgactcgcat	gttttccctg	accaaataccg	ccagaaggcg	ctcgccgccc	agcgatagca	6120
gttcttgcaa	ggaagcaaaag	tttttcaacg	gtttgagacc	gtccgcctga	ggcatgcttt	6180
tgagcgtttg	accaagcagt	tccaggcggg	cccacagctc	ggtcacctgc	tctacggcat	6240
ctcgatccag	catatctcct	cgtttcgcgg	gttggggcgg	ctttcgctgt	acggcagtag	6300
tcggtgctcg	tccagacggg	ccagggctcat	gtctttccac	gggcgcaggg	tcctcgtag	6360
cgtagtctgg	gtcacgggtga	aggggtgcgc	tccgggctgc	gcgctggcca	gggtgcgctt	6420
gaggctggtc	ctgctgggtgc	tgaagcgctg	ccggctcttcg	ccctgcgcgt	cggccaggta	6480
gcatttgacc	atgggtgtcat	agtccagccc	ctccgcggcg	tggcccttgg	cgcgagctt	6540
gcccttggag	gaggcgccgc	acgaggggca	gtgcagactt	ttgagggcgt	agagcttggg	6600
cgcgagaaat	accgattccg	gggagtaggc	atccgcgcgc	caggccccgc	agacggtctc	6660
gcattccacg	agccagggtga	gctctggccg	ttcgggggtca	aaaaccagggt	ttcccccatg	6720
ctttttgatg	cgtttctttac	ctctgggtttc	catgagccgg	tgtccacgct	cggtgacgaa	6780
aaggctgtcc	gtgtccccgt	atacagactt	gagaggcctg	tcctcgagcg	gtgttccgcg	6840
gtcctcctcg	tatagaaact	cggaccactc	tgagacaaaag	gctcgcgtcc	aggccagcac	6900
gaaggaggct	aagtgggagg	ggtagcggtc	gttgtccact	aggggggtcca	ctcgctccag	6960
ggtgtgaaga	cacatgtcgc	cctcttcggc	atcaaggaaag	gtgattgggt	tgtaggtgta	7020
ggccacgtga	ccgggtgttc	ctgaaggggg	gctataaaaag	ggggtggggg	cgcgcttcgc	7080
ctcactctct	tcgcgcatcg	tgtctgcgag	ggccagctgt	tggggtgagt	actccctctg	7140
aaaagcgggc	atgacttctg	cgctaagatt	gtcagtttcc	aaaaacgagg	aggattttgat	7200
attcacctgg	cccgcggtga	tgcctttgag	ggtggccgca	tccatctggt	cagaaaagac	7260
aatctttttg	ttgtcaagct	tgggtggcaaa	cgacccgtag	agggcggttg	acagcaactt	7320
ggcgtaggag	cgcagggttt	ggttttttgtc	gcgatcggcg	cgctccttgg	ccgcgatggt	7380
tagctgcacg	tattcgcgcg	caacgcaccg	ccattcggga	aagacggttg	tgcgctcgct	7440
gggcaccagg	tgcacgcgcc	aaccgcgggt	gtgcagggtg	acaagggtcaa	cgctgggtggc	7500
tacctctccg	cgtaggcgct	cgttgggtcca	gcagaggcgg	ccgcccttgc	gcgagcagaa	7560
tggcggtagg	gggtctagct	gcgtctcgtc	cgggggggtct	gcgtccacgg	taaagacccc	7620
gggcagcagg	cgcgcgtcga	agtagtctat	cttgcatcct	tgcaagtcta	gcgcctgctg	7680
ccatgcgcgg	gcggcaagcg	cgcgctcgta	tgggttgagt	gggggacccc	atggcatggg	7740
gtgggtgagc	gcggaggcgt	acatgccgca	aatgtcgtaa	acgtagaggg	gctctctgag	7800
tattccaaga	tatgtagggt	agcatcttcc	accgcggatg	ctggcgcgca	cgtaatcgta	7860
tagttcgtgc	gagggagcga	ggaggtcggg	accaggttg	ctacggggcg	gctgctctgc	7920
tcggaagact	atctgcctga	agatggcatg	tgggttgatg	gatattgggtg	gacgctggaa	7980
gacgttgaag	ctggcgctctg	tgagacctac	cgcgctcagc	acgaaggagg	cgtaggagtc	8040
gcgcagcttg	ttgaccagct	cggcggtgac	ctgcacgtct	agggcgagct	agtccagggt	8100
ttccttgatg	atgtcatact	tatcctgtcc	cttttttttct	cacagctcgc	ggttgaggac	8160
aaactcttcg	cggctctttcc	agtactcttg	gatcggaac	ccgtcggcct	ccgaacggta	8220
agagcctagc	atgtagaact	ggttgacggc	ctggtagggc	cagcatccct	tttctacggg	8280
tagcgcgtat	gcctgcgcgg	ccttccggag	cgagggtgtg	gtgagcgcaa	aggtgtccct	8340
gaccatgact	ttgaggtaact	ggtatttgaa	gtcagtgctg	tcgcatccgc	cctgctccca	8400
gagcaaaaaa	tccgtgcgct	ttttggaacg	cggattttggc	agggcggaagg	tgacatcggt	8460
gaagagtatc	tttcccgcgc	gaggcataaa	gttgctgtg	atgcggaagg	gtcccggcac	8520
ctcggaacgg	ttgttaatta	cctggggcggc	gagcagatc	tcgtcaaagc	cgttgatggt	8580
gtggcccaca	atgtaaagtt	ccaagaagcg	cgggatgccc	ttgatggaag	gcaatttttt	8640

-27-

aagttcctcg	taggtgagct	cttcaggsga	gctgagcccc	tgctctgaaa	gggcccagtc	8700
tgcaagatga	gggttggaag	cgacgaatga	gctccacagg	tcacgggcca	ttagcatttg	8760
caggtggtcg	cgaaagggtcc	taaactggcg	acctatggcc	attttttctg	gggtgatgca	8820
gtagaaggta	agcgggtctt	gttcccagcg	gtcccatcca	aggttcgcgg	ctaggtctcg	8880
cgcggcagtc	actagaggct	catctccgcc	gaacttcatg	accagcatga	agggcacgag	8940
ctgcttccca	aaggccccc	tccaagtata	ggtctctaca	tcgtagggtga	caaagagacg	9000
ctcgggtgcga	ggatgcgagc	cgatcgggaa	gaactggatc	tcccgccacc	aattggagga	9060
gtggctattg	atgtggtgaa	agtagaagtc	cctgcgacgg	gccgaacact	cgtgctggct	9120
tttgtaaaaa	cgtgcgcagt	actggcagcg	gtgcacgggc	tgtacatcct	gcacgaggtt	9180
gacctgacga	ccgcgcacaa	ggaagcagag	tgggaatttg	agccccctcg	ctggcgggtt	9240
tggctggtgg	tcttctactt	cggctgcttg	tccctgaccg	tctggctgct	cgaggggagt	9300
tacgggtggat	cggaccacca	cgcgcgcga	gccccaaagt	cagatgtccg	cgcgcggcgg	9360
tcggagcttg	atgacaacat	cgcgcagatg	ggagctgtcc	atggtctgga	gctcccgcgg	9420
cgtcaggtca	ggcgggagct	cctgcaggtt	tacctcgcat	agacgggtca	gggcgcgggc	9480
tagatccagg	tgatacctaa	tttccagggg	ctggttggtg	gcggcgtcga	tggcttgcaa	9540
gaggcccgat	ccccgcggcg	cgactacggt	accgcgcggc	gggcgggtggg	cgcggggggt	9600
gtccttggtg	gatgcatcta	aaagcgggtg	cgcgggcggg	cccccgagg	tagggggggc	9660
tccggacccg	ccgggagagg	gggcaggggc	acgtcggcgc	cgcgcgcggg	caggagctgg	9720
tgctgcgcgc	gtaggttgct	ggcgaacgcg	acgacgcggc	ggttgatctc	ctgaatctgg	9780
cgctctgcgc	tgaagacgac	gggcccgggt	agcttgagcc	tgaagagag	ttcgacagaa	9840
tcaatttcgg	tgtcgttgac	ggcggcctgg	cgcaaaatct	cctgcacgtc	tcctgagttg	9900
tcttgatagg	cgatctcggc	catgaactgc	tcgatctctt	cctcctggag	atctccgcgt	9960
ccggctcgct	ccacggtggc	ggcgaggtcg	ttggaaatgc	gggccatgag	ctgcgagaag	10020
gcgttgaggc	ctccctcggt	ccagacgcgg	ctgtagacca	cgcccccttc	ggcatcgcg	10080
gcgcgcagta	ccacctgcgc	gagattgagc	tccacgtgcc	gggcgaagac	ggcgtagttt	10140
cgcaggcgct	gaaagaggta	gttgagggtg	gtggcggtgt	gttctgccac	gaagaagtac	10200
ataaccagc	gtcgcaacgt	ggattcggtg	atatcccca	aggcctcaag	gcgctccatg	10260
gcctcgtaga	agtccacggc	gaagttgaaa	aactgggagt	tgcgcgccga	cacggttaac	10320
tcctcctcca	gaagacggat	gagctcggcg	acagtgctcg	gcacctcgcg	ctcaaaggct	10380
acagggcgct	cttcttcttc	ttcaatctcc	ttctccataa	gggcctcccc	ttcttcttct	10440
tctggcggcg	gtgggggagg	ggggacacgg	cggcgacgac	ggcgacccgg	gaggcggtcg	10500
acaaagcgct	cgatcatctc	cccgcggcga	cggcgcatgg	tctcggtgac	ggcgcgcccg	10560
ttctcgcggg	ggcgcagttg	gaagacgccg	cccgtcatgt	cccgtttatg	ggttggcggg	10620
gggctgccat	gcggcagggg	tacggcgcta	acgatgcata	tcaacaattg	ttgtgtaggt	10680
actccgcgc	cgagggaacct	gagcgagtc	cgactgacgg	gatcggaata	cctctcgaga	10740
aaggcgctcta	accagtcaca	gtcgcaagggt	aggctgagca	ccgtggcggg	cggcagcggg	10800
cggcggtcgg	ggttggtttct	ggcggagggtg	ctgctgatga	tgtaatataa	gtaggcggtc	10860
ttgagacggc	ggatggtcga	cagaagcacc	atgtccttgg	gtccggcctg	ctgaatgcgc	10920
aggcggtcgg	ccatgcccc	ggcttcggtt	tgacatcggc	gcaggtcttt	gtagtagtct	10980
tgcctagagc	tttctacggg	cacttctctt	tctccttctt	cttgtcctgc	atctcttgca	11040
tctatcgctg	cggcggcggc	ggagtttggt	cgtaggtggc	gccctcttcc	tcccatgcgt	11100
gtgaccccg	agccccctcat	cggctgaagc	agggctaggt	cggcgacaac	gcgctcggct	11160
aatatggcct	gctgcacctg	cgtgagggtg	gactggaagt	catccatgtc	cacaaagcgg	11220
tggtagccgc	ccgtgttgat	ggtgtaagtg	cagttggcca	taacggacca	gttaacgggtc	11280
tggtagccgc	gctgcgagag	ctcgggtgtac	ctgagacgcg	agtaagccct	cgagtcaaat	11340
acgtagtcgt	tgcaagtccg	caccaggtac	tggatcccca	ccaaaaagtg	cggcggcggc	11400
tggcggtaga	ggggccagcg	taggggtggc	ggggctccgg	gggcgagatc	ttccaacata	11460
aggcgatgat	atccgtagat	gtacctggac	atccaggtga	tgcggcgggc	ggtggtggag	11520
gcgcgcggaa	agtgcgcggac	gcggttccag	atggtgcgca	gcggcaaaaa	gtgctccatg	11580
gtcgggacgc	tctggccggg	caggcgcgcg	caatcggtga	cgctctagac	cgtgcaaaag	11640
gagagcctgt	aagcgggcac	tcttccgtgg	tctgggtggat	aaattcgcaa	gggtatcatg	11700
gcggacgacc	ggggttcgag	ccccgtatcc	ggcgcgtccg	cgtgatccat	gcggttaccg	11760
cccgcgtgct	gaaccagggt	gtgcgacgtc	agacaacggg	ggagtgtctc	ttttggcttc	11820
cttccaggcg	cgcggtctgc	tgcgtagctt	tgtttggcca	ctggccgcgc	gcagcgtaag	11880
cggttaggct	ggaaagcgaa	agcattaagt	ggctcgctcc	ctgtagccgg	agggttatatt	11940
tccaagggtt	gagtcgcggg	acccccgggt	cgagtctcgg	accggccgga	ctgcggcgaa	12000
cggggggttt	cctccccgtc	atgcaagacc	ccgcttgcaa	attcctccgg	aaacaggggac	12060
gagccccctt	tttgcttttc	ccagatgcac	ccggtgctgc	ggcagatgcg	ccccctctct	12120
cagcagcggc	aagacgaaga	gcagcggcag	acatgcaggg	caccctcccc	tcctcctacc	12180
gcgtcaggag	gggcgacatc	cgcgggtgac	gcggcagcag	atggtgatta	cgaacccccg	12240
cggcgccggg	cccggcacta	cctggacttg	gaggaggggc	agggcctggc	gcggctagga	12300

gcgcacctctc	ctgagcggta	cccaaggggtg	cagctgaagc	gtgatacgcg	tgaggcggtac	12360
gtgccgcggc	agaacctgtt	tccgcgaccgc	gagggagagg	agcccgagga	gatgcgggat	12420
cgaaagtccc	acgcagggcg	cgagctgcgg	catggcctga	atcgcgagcg	gttgctgcgc	12480
gaggaggact	ttgagcccga	cgcgcgaacc	gggattagtc	ccgcgcgcgc	acacgtggcg	12540
gccgcccgacc	tggttaaccgc	atacgagcag	acgggtgaacc	aggagattaa	ctttcaaaaa	12600
agctttaaca	accacgtgcg	tacgcttggtg	gcgcgcgagg	aggtggctat	aggactgatg	12660
catctgtggg	actttgttaag	cgcgctggag	caaaacccaa	atagcaagcc	gctcatggcg	12720
cagctgttcc	ttatagtgcg	gcacagcagg	gacaacgagg	cattcagggg	tgcgctgcta	12780
aacatagtag	agcccagagg	ccgctggctg	ctcgatttga	taaacatcct	gcagagcata	12840
gtggtgcagg	agcgcagcct	gagcctggct	gacaaggtgg	ccgccatcaa	ctattccatg	12900
cttagcctgg	gcaagtttta	cgcccgcgaag	atataccata	ccccttacgt	tcccatagac	12960
aaggaggtaa	agatcgaggg	gttctacatg	cgcatggcgc	tgaaggtgct	taccttgagc	13020
gacgacctgg	gcgtttatcg	caacgagcgc	atccacaagg	ccgtgagcgt	gagccggcgg	13080
cgcgagctca	gcgaccgcga	gctgatgcac	agcctgcaaa	gggcccctgg	tggcacggcg	13140
agcggcgata	gagaggccga	gtcctacttt	gacgcggggc	ctgacctgcg	ctgggccccca	13200
agccgacgcg	ccctggaggc	agctggggcc	ggacctgggc	tgggcggtgg	acccgcgcgc	13260
gctggcaacg	tcggcgggcgt	ggaggaatat	gacgaggacg	atgagtacga	gccagaggac	13320
ggcgagtact	aagcggtgat	gtttctgata	agatgatgca	agacgcaacg	gacccggcgg	13380
tgccggcgcg	gtgcagagc	cagccgtccg	gccttaactc	cacggacgac	tggcgccagg	13440
tcattggaccg	catcatgtcg	ctgactgcgc	gcaatcctga	cgcgttccgg	cagcagccgc	13500
aggccaaccg	gctctccgca	attctggaag	cggtgggtccc	ggcgcgcgca	aacccccacg	13560
acgagaaggt	gctggcgatc	gtaaacgcgc	tgcccgaaaa	cagggccatc	cggcccgacg	13620
aggccggcct	ggtctacgac	gcgctgcttc	agcgcgtggc	tcgttacaac	agcggcaacg	13680
tgcagaccaa	cctggaccgg	ctgggtggggg	atgtgcgcga	ggccgtggcg	cagcgtgagc	13740
gcgcgcagca	gcagggcaac	ctgggctcca	tggttgact	aaacgccttc	ctgagtacac	13800
agcccgccaa	cgtgccgcgg	ggacaggagg	actaccacaa	ctttgtgagc	gcactgcggc	13860
taatggtgac	tgagacaccg	caaagtggag	tgtaccagtc	tgggccagac	tattttttcc	13920
agaccagtag	acaaggcctg	cagaccgtaa	acctgagcca	ggctttcaaa	aacttgagg	13980
ggctgtgggg	ggtgcgggct	cccacaggcg	accgcgcgac	cgtgtctagc	ttgctgacgc	14040
ccaactcgcg	cctgttgctg	ctgctaatag	cgcccttcac	ggacagtggc	agcgtgtccc	14100
gggacacata	cctaggtcac	ttgctgacac	tgtaccgcga	ggccataggt	caggcgcgat	14160
tggacgagca	tactttccag	gagattacaa	gtgtcagccg	cgcgctgggg	caggaggaca	14220
cgggcagcct	ggaggcaacc	ctaaactacc	tgctgaccaa	ccggcggcag	aagatcccct	14280
cgttgacacg	tttaaacagc	gaggaggagc	gcattttgctg	ctacgtgcag	cagagcgtga	14340
gccttaacct	gatgcgcgac	ggggtaaacgc	ccagcgtggc	gctggacatg	accgcgcgca	14400
acatggaacc	gggcatgtat	gcctcaaacc	ggccgtttat	caaccgccta	atggactact	14460
tgcatcgcg	ggccgcctg	aaccccgagt	atttcaccaa	tgccatcttg	aacccgcact	14520
ggctaccgcc	ccctggtttc	tacaccgggg	gattcgaggt	gcccaggggt	aacgatggat	14580
tcctctggga	cgacatagac	gacagcgtgt	tttcccgcga	accgcagacc	ctgctagagt	14640
tgcaacagcg	cgagcaggca	gaggcggcgc	tgcgaaagga	aagcttccgc	aggccaagca	14700
gcttgctcga	tctaggcgct	gcggccccgc	ggtcagatgc	tagtagccca	tttccaaagt	14760
tgatagggtc	tcttaccagc	actcgcacca	cccgcccgcg	cctgctgggg	gaggaggagt	14820
acctaatacaa	ctcgctgctg	cagccgcagc	gcgaaaaaaa	cctgcctccg	gcatttccca	14880
acaacgggat	agagagccta	gtggacaaga	tgagtagatg	gaagacgtac	gcgcaggagc	14940
acagggacgt	gccaggccccg	cgccccccca	ccgctcgta	aaggcacgac	cgtcagcggg	15000
gtctggtgtg	ggaggacgat	gactcggcag	acgacagcag	cgtcctggat	ttggggaggga	15060
gtggcaaccc	gtttgcgcac	cttcgccccca	ggctggggag	aatgttttaa	aaaaaaaaaa	15120
gcatgatgca	aaataaaaaa	ctcaccaagg	ccatggcacc	gagcgttggt	tttcttgat	15180
tccccttagt	atgcggcgcg	cggcgatgta	tgaggaaggt	cctcctccct	cctacgagag	15240
tgtggtgagc	gcggcgccag	tgccggcgccg	gctgggttct	cccttcgatg	ctcccctgga	15300
cccggctttt	gtgcctccgc	ggtacctgcg	gcctaccggg	gggagaaaca	gcatccgtta	15360
ctctgagttg	gcacccctat	tcgacaccac	ccgtgtgtac	ctggtggaca	acaagtcaac	15420
ggatgtggca	tccctgaact	accagaacga	ccacagcaac	tttctgacca	cggtcattca	15480
aaacaatgac	tacagcccgg	gggaggcaag	cacacagacc	atcaatcttg	acgaccggtc	15540
gcactggggc	ggcgacctga	aaaccatcct	gcatacagac	atgccaaatg	tgaacgagtt	15600
catgtttacc	aataagttta	aggcgggggt	gatgggtgtc	cgcttgccca	ctaaggacaa	15660
tcaggtggag	ctgaaatacg	agtgggtgga	gttcacgctg	cccaggggca	actactccga	15720
gaccatgacc	atagacctta	tgaacaacgc	gatcgtggag	cactacttga	aagtgggcag	15780
acagaacggg	gttctggaaa	gcgacatcgg	ggtaaggttt	gacacccgca	acttcagact	15840
gggggttgac	cccgctcact	gtcttgctat	gcctggggta	tatacaaacg	aagccttcca	15900
tccagacatc	atthttgctgc	caggatgcgg	gggtggacttc	acccacagcc	gcctgagcaa	15960

-29-

cttggttgggc	atccgcaagc	ggcaaccctt	ccaggagggc	tttaggatca	cctacgatga	16020
tctggagggg	ggtaacattc	ccgcactggt	ggatgtggac	gcctaccagg	cgagcttgaa	16080
agatgacacc	gaacagggcg	ggggtggcgc	agggcgccagc	aacagcagtg	gcagcggcgc	16140
ggaagagaa	tccaacgcgg	cagccgcggc	aatgcagccg	gtggaggaca	tgaacgatca	16200
tgccattcgc	ggcgacacct	ttgccacacg	ggctgaggag	aagcgcgctg	aggccgaagc	16260
agcggccgaa	gctgccgccc	ccgctgcgca	acccgaggtc	gagaagcctc	agaagaaacc	16320
ggtgatcaaa	ccccgacag	aggacagcaa	gaaacgcagt	tacaacctaa	taagcaatga	16380
cagcaccttc	accagtagcc	gcagctggta	ccttgcatat	aactacggcg	accctcagac	16440
cggaatccgc	tcatggaccc	tgctttgcac	tcctgacgta	acctgcggct	cggagcaggt	16500
ctactggtcg	ttgccagaca	tgatgcaaga	ccccgtgacc	ttccgctcca	cgcgccagat	16560
cagcaacttt	ccggtggtgg	gcgccgagct	gttgcccgtg	cactccaaga	gcttctacaa	16620
cgaccaggcc	gtctactccc	aactcatccg	ccagtttacc	tctctgaccc	acgtgttcaa	16680
tcgctttccc	gagaaccaga	ttttggcgcg	cccgccagcc	cccaccatca	ccaccgtcag	16740
tgaaaacggt	cctgctctca	cagatcacgg	gacgctaccg	ctgcgcaaca	gcatcggagg	16800
agtccagcga	gtgaccatta	ctgacgccag	acgcccaccc	tgcccctacg	tttacaagcc	16860
cctgggcata	gtctcgccgc	gcgtcctatc	gagccgcact	ttttgagcaa	gcatgtccat	16920
ccttatatcg	cccagcaata	acacaggctg	gggcctgcgc	ttcccaagca	agatgtttgg	16980
cgggggccaag	aagcgtctccg	accaacaccc	agtgcgcgtg	cgcggggcact	accgcgcgcc	17040
ctggggcgcg	cacaaacgcg	gccgcactgg	gcgcaccacc	gtcgaatgacg	ccatcgacgc	17100
ggtggtggag	gaggcgcgca	actacacgcc	cacgcgcgca	ccagtgtcca	cagtggacgc	17160
ggccattcag	accgtggtgc	gcggagcccg	gcgctatgct	aaaatgaaga	gacggcgagg	17220
gcgcgtagca	cgtcgccacc	gccgcgcgac	cggcactgcc	gcccacgcgc	cggcggcggc	17280
cctgcttaac	cgcgcacgct	gcaccggccg	acgggcggcc	atgcggggccg	ctcgaaggct	17340
ggccgcgggt	attgtcactg	tgccccccag	gtccaggcga	cgagcggccg	ccgcagcagc	17400
cgcggccatt	agtgtctatg	ctcagggtcg	caggggccaac	gtgtattggg	tgccgcgactc	17460
ggttagcgcc	ctgcgcgtgc	ccgtgcgcac	ccggcccccg	cgcaactaga	ttgcaagaaa	17520
aaactactta	gactcgtact	gttgtatgta	tcacgcggcg	gcggcgcgca	acgaagctat	17580
gtccaagcgc	aaaatcaaa	aagagatgct	ccaggctcat	gcgcgggaga	tctatggccc	17640
cccgaagaag	gaagagcagg	attacaagcc	ccgaaagcta	aagcgggtca	aaaagaaaaa	17700
gaaagatgat	gatgatgaac	ttgacgacga	ggtggaactg	ctgcacgcta	ccgcgcgccag	17760
gcgacgggta	cagtggaaa	gtcgacgcgt	aaaacgtggt	ttgcgacccg	gcaccaccgt	17820
agtcttttac	cccgttgagc	gctccacccg	cacctacaag	cgcgtgtatg	atgaggtgta	17880
cggcgacgag	gacctgcttg	agcaggccaa	cgagcgcctc	ggggagtttg	cctacggaaa	17940
gcggcataag	gacatgctgg	cgttgccgct	ggacgagggc	aacccaacac	ctagcctaaa	18000
gcccgtaaac	ctgcagcagg	tgctgcccgc	gcttgccaccg	tccgaagaaa	agcgcggcct	18060
aaagcgcgag	tctggtgact	tggcaccac	cgtgcagctg	atggtaccca	agcgcgcagc	18120
actggaagat	gtcttgga	aaatgaccgt	ggaacctggg	ctggagcccc	aggtccgcgt	18180
gcggccaatc	aagcaggtgg	cgccgggact	gggcgtgacg	accgtggacg	ttcagatacc	18240
cactaccagt	agcaccagta	ttgccaccgc	cacagaggcg	atggagacac	aaacgtcccc	18300
ggttgcctca	gcggtggcgg	atgcgcgggt	gcaggcggtc	gctgcggccg	cgtccaagac	18360
ctctacggag	gtgcaaacgg	acccgtggat	gtttcgcgtt	tcagcccccc	ggcgcccgcg	18420
cggttcgagg	aagtacggcg	ccgccagcgc	gctactgccc	gaatatgccc	tacatccttc	18480
cattgcgcct	acccccggct	atcgtggcta	cacctaccgc	cccagaagac	gagcaactac	18540
ccgacgccga	accaccactg	gaacccgcgc	ccgcgctcgc	cgtcgccagc	ccgtgctggc	18600
cccgatattc	gtgcgcaggg	tggctcgcga	aggaggcagg	accctggtgc	tgccaacagc	18660
gcgctaccac	cccagcatcg	tttaaaagcc	ggtctttgtg	gttcttgcag	atatggccct	18720
cacctgccgc	ctccgtttcc	cgggtgccggg	attccgagga	agaatgcacc	gtaggagggg	18780
catggcccg	cacggcctga	cgggcggcat	gcgtcgtgcg	caccaccggc	ggcgccgcgc	18840
gtcgacccgt	cgcattgcgcg	gcgggtatcct	gccccctcct	attocactga	tcgcccgcgc	18900
gattggcgcc	gtgcccggaa	ttgcatccgt	ggccttgacg	gcgcagagac	actgattaaa	18960
aacaagtgtg	atgtggaaaa	atcaaaaata	aaagtctgga	ctctcacgct	cgttgggtcc	19020
tgtaactatt	ttgtagaatg	gaagacatca	actttgcgtc	tctggccccg	cgacacggct	19080
cgcgcccgtt	catgggaaac	tggcaagata	ctgggaccag	caatatgagc	ggtggcgcc	19140
tcagctgggg	ctcgctgtgg	agcggcatta	aaaatttcgg	ttccaccgtt	aagaactatg	19200
gcagcaaggc	ctggaacagc	agcacaggcc	agatgctgag	ggataagttg	aaagagcaaa	19260
atttccaaca	aaaggtggta	gatggcctgg	cctctggcat	tagcgggggtg	gtggacctgg	19320
ccaaccaggc	agtgcaaaa	aagattaaca	gtaagcttga	tccccgcctt	cccgtagagg	19380
agcctccacc	ggcgtggag	acagtgtctc	acagggggcg	tggcgaaaag	cgtccgcgcc	19440
ccgacaggga	agaaactctg	gtgacgcaaa	tagacgagcc	tcctctgtac	gaggaggcac	19500
taaagcaagg	cctgcccacc	acccgtccca	tcgcgcccac	ggctaccgga	gtgctgggccc	19560
agcacacacc	cgtaacgctg	gacctgcctc	ccccgcgcca	caccacgacg	aaacctgtgc	19620

-30-

tgccaggccc	gaccgcggtt	gttgtaaccc	gtcctagccg	cgcgtccctg	cgccgcgccc	19680
ccagcgggtcc	gcatcggtt	cggcccgtag	ccagtgggcaa	ctggcaaagc	acactgaaca	19740
gcatcggtggg	tctgggggtg	caatccctga	agcgccgacg	atgcttctga	atagctaacc	19800
tgtcgatatgt	gtgtcatgta	tgcgtccatg	tcgcccgcag	aggagctgct	gagccgccc	19860
gcgcccgcctt	tccaagatgg	ctaccccttc	gatgatgccg	cagtggctctt	acatgcacat	19920
ctcggggccag	gacgcctcgg	agtacctgag	ccccgggctg	gtgcagtttg	cccgcgccac	19980
cgagacgtac	ttcagcctga	ataacaagtt	tagaaacccc	acggtggcgc	ctacgcacga	20040
cgtgaccaca	gaccggtccc	agcgtttgac	gctgcggttc	atccctgtgg	accgtgagga	20100
tactgcgtac	tcgtacaagg	cgcggttcac	cctagctgtg	ggtgataacc	gtgtgctgga	20160
catggcttcc	acgtactttg	acatccgcgg	cgtgctggac	agggggcccta	cttttaagcc	20220
ctactctggc	actgcctaca	acgcccgtgg	tcccaagggt	gccccaaatc	cttgcgatg	20280
ggatgaagct	gctactgctc	ttgaaataaa	cctagaagaa	gaggacgatg	acaacgaaga	20340
cgaagtagac	gagcaagctg	agcagcaaaa	aactcacgta	tttgggcagg	cgccttatcc	20400
tgggtataaat	attacaaagg	agggtattca	aataggtgtc	gaaggtcaaa	cacctaaata	20460
tgccgataaaa	acattttcaac	ctgaacctca	aataggagaa	tctcagtggg	acgaaactga	20520
aattaatcat	gcagctggga	gagtccttaa	aaagactacc	ccaatgaaac	catgttaacg	20580
ttcatatgca	aaacccacaa	atgaaaatgg	agggcaaggc	attcttgtaa	agcaacaaaa	20640
tggaaagcta	gaaagtcaag	tggaaatgca	attttttctca	actactgagg	cgaccgcagg	20700
caatggtgat	aacttgactc	ctaaagtggg	attgtacagt	gaagatgtag	atatagaaac	20760
cccagacact	catattttctt	acatgcccac	tattaaggaa	ggtaaactcac	gagaactaat	20820
gggccaacaa	tctatgocca	acaggcctaa	ttacatttgc	tttagggaca	attttattgg	20880
tctaattgtat	tacaacagca	cgggtaatat	gggtgttctg	gcggggccaag	catcgcagtt	20940
gaatgctggt	gtagatttgc	aagacagaaa	cacagagctt	tcataccagc	ttttgcttga	21000
ttccattggt	gatagaacca	ggtacttttc	tatgtggaat	caggctgttg	acagctatga	21060
tccagatggt	agaattattg	aaaatcatgg	aactgaagat	gaacttccaa	attactgctt	21120
tccactggga	ggtgtgatta	atacagagac	tcttaccaaag	gtaaaaccta	aaacagggtca	21180
ggaaaatgga	tgggaaaaag	atgctacaga	attttcagat	aaaaatgaaa	taagagttgg	21240
aaataatttt	gccatggaaa	tcaatctaaa	tgccaacctg	tggagaaatt	tcctgtactc	21300
caacatagcg	ctgtatttgc	ccgacaagct	aaagtacagt	ccttccaacg	taaaaatttc	21360
tgataaccca	aacacctacg	actacatgaa	caagcgagtg	gtggctccc	ggttagtgga	21420
ctgctacatt	aaccttggag	cacgctggtc	ccttgactat	atggacaacg	tcaaccatt	21480
taaccaccac	cgcaatgctg	gcctgcgcta	ccgctcaatg	ttgctgggca	atggtcgcta	21540
tgtgcccttc	cacatccagg	tgctcagaa	gttctttgcc	attaaaaacc	tccttctcct	21600
gccgggctca	tacacctacg	agtggaaact	caggaaggat	gttaacatgg	ttctgcagag	21660
ctccctagga	aatgacctaa	gggttgacgg	agccagcatt	aagtttgata	gcatttgcct	21720
ttacgccacc	ttcttcccca	tggcccacaa	caccgcctcc	acgcttgagg	ccatgcttag	21780
aaacgacacc	aacgaccagt	cctttaacga	ctatctctcc	gccgccaaca	tgctctaccc	21840
tatacccgcc	aacgctacca	acgtgcccac	atccatcccc	tcccgcaact	gggcggtttt	21900
ccgcggcttg	gccttcacg	gccttaagac	taaggaaacc	ccatcactgg	gctcgggcta	21960
cgaccccttat	tacacactct	ctggctctat	accctaccta	gatggaacct	tttacctcaa	22020
ccacaccttt	aagaagggtg	ccattacctt	tgactcttct	gtcagctggc	ctggcaatga	22080
ccgcctgctt	acccccaacg	agtttgaaat	taagcgctca	gttgacgggg	agggttacaa	22140
cgttgcccg	tgtaacatga	ccaaagactg	gttccctggta	caaatgctag	ctaactacaa	22200
cattggctac	cagggtctct	atatcccaga	gagctacaag	gaccgcatgt	actccttctt	22260
tagaaaacttc	cagcccatga	gccgtcaggt	ggtggatgat	actaaatata	aggactacca	22320
acaggtgggc	atcctacacc	aacacaacaa	ctctggattt	gttggctacc	ttgccccac	22380
catgcgcgaa	ggacaggcct	accctgctaa	cttcccctat	ccgcttatag	gcaagaccgc	22440
agttgacagc	attacccaga	aaaagtttct	tgcgactcgc	accctttggc	gcacccatt	22500
ctccagtaac	tttatgtcca	tgggcgcact	cacagactcg	ggccaaaacc	ttctctacgc	22560
caactccgcc	cacgcgctag	acatgacttt	tgagggtggat	cccatggacg	agccccacct	22620
tctttatggt	ttgtttgaag	tctttgacgt	ggtccgtgtg	caccggccgc	accgcggcgt	22680
catcgaaacc	gtgtacctgc	gcacgcctt	ctcggccggc	aacgccacaa	cataaagaag	22740
caagcaacat	caacaacagc	tgccgccatg	ggctccagtg	agcaggaact	gaaagccatt	22800
gtcaaagatc	ttggttgtgg	gccatatttt	ttgggcacct	atgacaagcg	ctttccaggc	22860
tttgtttctc	cacacaagct	cgctgcgc	atagtcaata	cggccggtcg	cgagactggg	22920
ggcgtaact	ggatggcctt	tgcttggaac	ccgcactcaa	aaacatgcta	cctctttgag	22980
ccctttggct	tttctgacca	gcgactcaag	caggtttacc	agtttgagta	cgagtcactc	23040
ctgcgcgta	gcgccattgc	ttcttccccc	gaccgctgta	taacgctgga	aaagtccacc	23100
caaagcgtac	aggggcccac	ctcgccgc	tgtggactat	tctgctgcat	gtttctccac	23160
gcctttgcca	actggcccca	aactcccatg	gatcacacac	ccaccatgaa	ccttattacc	23220
ggggtaacca	actccatgct	caacagtc	caggtacagc	ccaccctgcg	tcgcaaccag	23280

-31-

gaacagctct	acagcttctt	ggagcgccac	tgcgcctact	tccgcagcca	cagtgcgcag	23340
attaggagcg	ccacttcttt	ttgtcacttg	aaaaacatgt	aaaaataatg	tacttagagac	23400
acttttcaata	aaggcaaatg	ctttttatttg	tacactctcg	ggtgattatt	tacccccacc	23460
cttgccgtct	gcgccgtttg	gggaggcggc	ggcgacgggg	acggggacga	cacgtcctcc	23520
atgggtgggg	gacgtcgcg	cgcaccgcgt	ccgcgctcgg	gggtggtttc	gcgctgctcc	23580
tcttcccgac	tggccatttc	cttctctctat	aggcagaaaa	agatcatgga	gtcagtcgag	23640
aagaaggaca	gcctaaccgc	cccctctgag	ttcgccacca	ccgcctccac	cgatgccgcc	23700
aacgcgccta	ccaccttccc	cgtcgaggca	cccccgcttg	aggaggagga	agtgattatc	23760
gagcaggacc	caggttttgt	aagcgaagac	gacgaggacc	gctcagtacc	aacagaggat	23820
aaaaagcaag	accaggacaa	cgcagaggca	aacgaggaac	aagtcgggcg	gggggacgaa	23880
aggcatggcg	actacctaga	tgtggggagac	gacgtgctgt	tgaagcatct	gcagcgccag	23940
tgcgccatta	tctgcgacgc	gttgcaagag	cgcagcgatg	tgcccctcgc	catagcggat	24000
gtcagccttg	cctacgaacg	ccacctatct	tcaccgcgcg	taccccccaa	acgccaagaa	24060
aacggcacat	gcgagcccaa	cccgcgcctc	aacttctacc	ccgtatttgc	cgtgccagag	24120
gtgcttgcca	cctatcacat	ctttttccaa	aactgcaaga	tacccctatc	ctgccgtgcc	24180
aaccgcagcc	gagcggacaa	gcagctggcc	ttgcggcagg	gcgctgtcat	acctgatatc	24240
gcctcgctca	acgaagtgcc	aaaaatcttt	gagggctctg	gacgcgacga	gaagcgcgcg	24300
gcaaacgctc	tgcaacagga	aaacagcgaa	aatgaaagtc	actctggagt	gttgggtggaa	24360
ctcgagggtg	acaacgcgcg	cctagccgta	ctaaaacgca	gcatcgaggt	caccactttt	24420
gcctaccggg	cacttaacct	accccccaag	gtcatgagca	cagtcatgag	tgagctgac	24480
gtgcgcctg	cgcagcccct	ggagagggat	gcaaatgtgc	aagaacaaac	agaggagggc	24540
ctaccgcgag	ttggcgacga	gcagctagcg	cgttggtctc	aaacgcgcga	gcctgccgac	24600
ttggaggagc	gacgcaaact	aatgatggcc	gcagtgctcg	ttaccgtgga	gcttgagtgc	24660
atgcagcgg	tctttgctga	cccgagatg	cagcgcaagc	tagaggaaac	attgcactac	24720
acctttcgac	agggtacgt	acgccaggcc	tgcaagatct	ccaacgtgga	gctctgcaac	24780
ctggtctcct	accttggaat	tttgacgaa	aaaccgcttg	ggcaaaacgt	gcttcattcc	24840
acgctcaagg	gcgaggcgcg	ccgcgactac	gtccgcgact	gcgtttactt	atctctatgc	24900
tacacctggc	agacggccat	gggcgttttg	cagcagtgct	tgaggagagt	caacctcaag	24960
gagctgcaga	aactgctaaa	gcaaaaactg	aaggacctat	ggacggcctt	caacgagcgc	25020
tccgtggccg	cgcacctggc	ggacatcatt	ttccccgaac	gcctgcttaa	aaccctgcaa	25080
cagggtctcg	cagacttcac	cagtcaaagc	atggtgcaga	actttaggaa	ctttatccta	25140
gagcgctcag	gaatcttgcc	cgccacctgc	tgtgcacttc	ctagcgactt	tgtgcccatt	25200
aagtaccgcg	aatgccctcc	gccgcttttg	ggccactgct	accttctgca	gctagccaac	25260
taccttgctt	accactctga	cataattgaa	gcagtgagcg	gtgacggtct	actggagtct	25320
cactgtcgct	gcaacctatg	caccccgcac	cgctccctgg	tttgcaattc	gcagctgctt	25380
aacgaaagtc	aaattatcgg	tacctttgag	ctgcagggtc	cctcgccctga	cgaaaagtcc	25440
gcggctccgg	ggttgaaact	cactccgggg	ctgtggacgt	cggcttacct	tcgcaaattt	25500
gtacctgagg	actaccacgc	ccacgagatt	aggttctacg	aagaccaatc	ccgcccgcga	25560
aatgcggagc	ttacgcctcg	cgtcattacc	cagggccaca	ttcttgggcca	attgcaagcc	25620
atcaacaaag	cccggcaaga	gtttctgcta	cgaaagggac	gggggggttta	cttggaacccc	25680
cagtcgggcg	aggagctcaa	cccaatcccc	ccgcccgcgc	agccctatca	gcagcagccg	25740
cgggcccctg	cttcccagga	tggcacccaa	aaagaagctg	cagctgccgc	cgccacccac	25800
ggacgaggag	gaatactggg	acagttaggc	agaggaggtt	ttggacgagg	aggaggagga	25860
catgatggaa	gactgggaga	gcctagacga	ggaagcttcc	gaggtcgaag	aggtgtcaga	25920
cgaaacaccg	tcacctcgg	tcgcattccc	ctgcgcggcg	ccccagaaat	cggaacccgg	25980
ttccagcatg	gctacaacct	ccgctcctca	ggcgccgcgc	gcaactgccc	ttcgccgacc	26040
caaccgtaga	tgggacacca	ctggaaccag	ggccggtaag	tccaagcagc	cgccgcctgt	26100
agcccaagag	caacaacagc	gccaaggcta	ccgctcatgg	cgcgggcaca	agaacgccat	26160
agttgcttgc	ttgcaagact	gtggggggcaa	catctccttc	gcccgcgcgt	ttcttctcta	26220
ccatcacggc	gtggccttcc	cccgtaacat	cctgcattac	taccgtcatc	tctacagccc	26280
atactgcacc	ggcggcagcg	gcagcggcag	caacagcagc	ggccacacag	aagcaaaggc	26340
gaccggatag	caagactctg	acaaagccca	agaaatccac	agcggcggca	gcagcaggag	26400
gaggagcgct	gcgtctggcg	cccaacgaa	ccgtatcgac	ccgcgagctt	agaaacagga	26460
tttttccac	tctgtatgct	atatttcaac	agagcagggg	ccaagaacaa	gagctgaaaa	26520
taaaaaacag	gtctctgcga	tccctcacc	gcagctgcct	gtatcacaaa	agcgaagatc	26580
agcttcggcg	cacgctggaa	gacgcggagg	ctctcttcag	taaatactgc	gcgctgactc	26640
ttaaggacta	gtttcgcgcc	ctttctcaaa	tttaagcgcg	aaaactacgt	catctccagc	26700
ggccacaccc	ggcgcgacga	cctgtcgta	ggccatttat	gagcaaggaa	attcccacgc	26760
cctacatgtg	gagttaccag	ccacaaatgg	gacttgcggc	tggagctgcc	caagactact	26820
caaccogaat	aaactacatg	agcgcgggac	cccacatgat	atcccgggtc	aacggaatcc	26880
gcgcccaccg	aaaccgaatt	ctcttggaac	aggcggtat	taccaccaca	cctcgtaata	26940

-32-

accttaatcc	ccgtagttgg	cccgtgccc	tggtgtacca	ggaaagtccc	gctccccacca	27000
ctgtggtact	tcccagagac	gcccaggccg	aagttcagat	gactaactca	ggggcgagc	27060
ttgcgggagg	ctttcgtcac	agggtgcggt	cgcccgggca	gggtataact	cacctgacaa	27120
tcagagggcg	aggtattcag	ctcaacgacg	agtcgggtgag	ctcctcgctt	ggtctccgtc	27180
cggacgggac	atttcagatc	ggcggcgccg	gocgtocctt	attcacgcct	cgtaggcaa	27240
tcctaactct	gcagacctcg	tcctctgagc	cgcgctctgg	aggcattgga	actctgcaat	27300
ttattgagga	gtttgtgcca	tcgggtctact	ttaacccctt	ctcgggacct	cccggccact	27360
atccggatca	atattattcct	aactttgacg	cggtaaaggga	ctcggcggac	ggctacgact	27420
gaatgttaag	tggagaggca	gagcaactgc	gcctgaaaca	cctggtccac	tgtcgccgcc	27480
acaagtgcct	tgcccgcgac	tcgggtgagt	tttgctactt	tgaattgccc	gaggatcata	27540
tcgagggccc	ggcgcacggc	gtccggctta	ccgcccggga	agagcttgcc	cgtagcctga	27600
ttcgggagtt	taccagcgcc	ccccgtctag	ttgagcggga	caggggaccc	tgtgttctca	27660
ctgtgatttg	caactgtcct	aaccttggat	tacatcaaga	tctttgttgc	catctctgtg	27720
ctgagtataa	taaatacaga	aattaaaata	tactggggct	cctatcgcca	tcctgtaaac	27780
gccaccgtct	tcacccgccc	aagcaaacca	aggcgaacct	tacctggtac	ttttaacatc	27840
tctccctctg	tgattttacaa	cagtttcaac	ccagacggag	tgagtctacg	agagaacctc	27900
tcagagctca	gtactcccat	cagaaaaaac	accaccctcc	ttacctgccg	ggaacgtacg	27960
agtgcgctcac	cggccgctgc	accacaccta	ccgcctgacc	gtaaaccaga	ctttttccgg	28020
acagacctca	ataactctgt	ttaccagaac	aggaggtgag	cttagaaaac	ccttagggta	28080
ttaggccaaa	ggcgcagcta	ctgtgggggt	tatgaacaat	tcaagcaact	ctacgggcta	28140
ttctaattca	ggtttctcta	gaaatggacg	gaattattac	agagcagcgc	ctgctagaaa	28200
gacgcagggc	agcggccgag	caacagcgca	tgaatcaaga	gctccaagac	atggttaact	28260
tgcaccagtg	caaaaagggg	atcttttgtc	tggtaaagca	ggccaaagtc	acctacgaca	28320
gtaataccac	cggacaccgc	cttagctaca	agttgccaac	caagcgtcag	aaattggtgg	28380
tcattggtggg	agaaaagccc	attaccataa	ctcagcactc	ggtagaaacc	gaaggctgca	28440
ttcactcacc	ttgtcaagga	cctgaggatc	tctgcacctc	tattaagacc	ctgtgcggtc	28500
tcaaagatct	tattcccttt	aactaataaa	aaaaaataat	aaagcatcac	ttacttaaaa	28560
tcagttagca	aattttctgtc	cagtttatct	agcagcacct	ccttgccctc	ctcccagctc	28620
tggtattgca	gcttccctct	ggctgcaaac	tttctccaca	atctaaatgg	aatgtcagtt	28680
tcctcctggt	ctgtccatc	cgcacccact	atcttctatg	tgttgagat	gaagcgcgca	28740
agaccgtctg	aagatacctt	caaccccggt	tatccatag	acacggaaac	cggtcctcca	28800
actgtgcctt	ttcttactcc	tccttttgta	tcctccaatg	ggtttcaaga	gagtccccct	28860
ggggtactct	ctttgcgctc	atccgaacct	ctagttacct	ccaatggcat	gcttgcgctc	28920
aaaatgggca	acggccctct	tctggacgag	gcccggcaacc	ttacctccca	aaatgtaacc	28980
actgtgagcc	cacctctcaa	aaaaaccaag	tcaaacataa	acctggaaat	atctgcaccc	29040
ctcacagtta	cctcagaagc	cctaactgtg	gctgcgcgcg	cacctctaata	ggtcgcgggc	29100
aacacactca	ccatgcaatc	acaggccccg	ctaaccgtgc	acgactccaa	acttagcatt	29160
gccacccaag	gacccctcac	agtgctcagaa	ggaaagctag	ccctgcaaac	atcaggcccc	29220
ctcaccacca	cggatagcag	tacccttact	atcactgcct	cacccctctc	aactactgcc	29280
actggtagct	ttggcattga	cttgaaagag	cccatttata	cacaaaatgg	aaaactagga	29340
ctaaagtacg	gggctccttt	gcatgtaaca	gacgacctaa	acactttgac	cgtagcaact	29400
gggtccaggtg	tgactattaa	taatacttcc	ttgcaaacata	aagttactgg	agccttgggt	29460
tttgattcac	aaggcaatat	gcaacttaata	gtagcaggag	gactaaggat	tgattctcaa	29520
aacagacgcc	ttataacttga	tgtagtttat	ccgtttgatg	ctcaaaaacca	actaaatcta	29580
agactaggac	agggccctct	ttttataaac	tcagcccaca	acttgatat	taactacaac	29640
aaaggccctt	acttggtttac	agcttcaaac	aattccaaaa	agcttgaggt	taacctagc	29700
actgccaaag	ggttgatgtt	tgacgctaca	gccatagcca	ttaatgcagg	agatgggctt	29760
gaatttggtt	cacctaatgc	accaaacaca	aatccctca	aaacaaaaat	tgcccatggc	29820
ctagaatttg	attcaaacaa	ggctatgggt	cctaaactag	gaactggcct	tagttttgac	29880
agcacaggtg	ccattacagt	aggaaacaaa	aataatgata	agctaacttt	gtggaccaca	29940
ccagctccat	ctcctaactg	tagactaaat	gcagagaaag	atgctaaact	cactttgggt	30000
ttaacaaaat	gtggcagcta	aatacttgct	acagtttcag	ttttggctgt	taaaggcagt	30060
ttggctccaa	tatctggaac	agttcaaagt	gctcatctta	ttataagatt	tgacgaaaat	30120
ggagtgtac	taaacaaatc	cttctggac	ccagaatatt	ggaactttag	aaatggagat	30180
cttactgaag	gcacagccta	tacaaacgct	gttggtattta	tgcttaacct	atcagcttat	30240
ccaaaatctc	acggtaaaac	tgccaaaagt	aacattgtca	gtcaagttta	cttaaacgga	30300
gacaaaacta	aacctgtaac	actaaccatt	acactaaaacg	gtacacagga	aacaggagac	30360
acaactccaa	gtgcatactc	tatgtcattt	tcattgggact	gggtctggcca	caactacatt	30420
aatgaaatat	ttgccacatc	ctcttacact	ttttcataca	ttgcccaga	ataaagaatc	30480
gtttgtgtta	tgtttcaacg	tgtttatatt	tcaattgcag	aaaatttcaa	gtcatttttc	30540
attcagtagt	atagcccccac	caccacatag	cttatacaga	tcaccgtacc	ttaatcaaac	30600

-33-

tcacagaacc	ctagtattca	acctgccacc	tcctcccaa	cacacagagt	acacagtcct	30660
ttctccccgg	ctggccttaa	aaagcatcat	atcatgggta	acagacatat	tcttaggtgt	30720
tatatccac	acggtttctt	gtcgagccaa	acgctcatca	gtgatattaa	taaactcccc	30780
gggcagctca	cttaagttca	tgctgctgtc	cagctgctga	gccacagggt	gctgtccaac	30840
ttgcggttgc	ttaacgggcg	gcgaaggaga	agtccacgcc	tacatggggg	tagagtcata	30900
atcgtgcata	aggatagggc	gggtggtgctg	cagcagcgcg	cgaataaaact	gctgccgccc	30960
ccgctccgctc	ctgcaggaat	acaacatggc	agtggtctcc	tcagcgatga	ttcgaccgcg	31020
ccgcagcata	aggcgcttgc	tcctccgggc	acagcagcgc	accctgatct	cacttaaact	31080
agcacagtaa	ctgcagcaca	gcaccacaat	attgttcaaa	atcccacagt	gcaaggcgct	31140
gtatccaaag	ctcatggcgg	ggaccacaga	acccacgtgg	ccatcatacc	acaagcgccg	31200
gtagattaag	tggcgacccc	tcataaacac	gctggacata	aacattacct	cttttggcat	31260
gttgtaattc	accacctccc	ggtagcatat	aaacctctga	ttaaactatgg	cgccatccac	31320
caccatccta	aaccagctgg	ccaaaacctg	cccgccggct	atacactgca	gggaaccggg	31380
actggaacaa	tgacagtggg	gagcccagga	ctcgtaacca	tggtatcatca	tgctcgatcat	31440
gatatacaat	tgggcacaa	acaggcacac	gtgcatacac	ttcctcagga	ttacaagctc	31500
ctcccgcggt	agaaccatat	cccagggaac	aacccttccc	tgaatcagcg	taaatcccac	31560
actgcaggga	agacctcgca	cgtaactcac	gttggtgcatt	gtcaaagtgt	tacattcggg	31620
cagcagcgga	tgatcctcca	gtatggtagc	gcgggtttct	gtctcaaaag	gaggtagacg	31680
atccctactg	tacggagtgc	gcccagacaa	ccgagatcgt	gttggtcgta	gtgtcatgcc	31740
aaatggaacg	ccggacgtag	tcataatttc	tgaagcaaaa	ccaggtgcgg	gcgtgacaaa	31800
cagatctgcg	tctccggtct	cgccgcttag	atcgctctgt	gtagtagttg	tagtatatcc	31860
actctctcaa	agcatccagg	cgccccctgg	cttcgggttc	tatgtaaaact	ccttcatgcg	31920
ccgctgccct	gataacatcc	accaccgcag	aataagccac	acccagccaa	cctacacatt	31980
cgttctgcga	gtcacacacg	ggaggagcgg	gaagagctgg	aagaaccatg	tttttttttt	32040
tattccaaaa	gattatccaa	aacctcaaaa	tgaagatcta	ttaagtgaac	gcgctcccct	32100
ccggtggcgt	gggtcaaaact	tacagccaaa	gaacagataa	tggtatttgt	aagatgttgc	32160
acaatggctt	ccaaaaggca	aacggccctc	acgtccaagt	ggacgtaaag	gctaaaccct	32220
tcagggtgaa	tctcctctat	aaacattcca	gcaccttcaa	ccatgcccaa	ataattctca	32280
tctcgccacc	ttctcaatat	atctctaagc	aaatcccga	tattaagtcc	ggccattgta	32340
aaaatctgct	ccagagcgcc	ctccaccttc	agcctcaagc	agcgaatcat	gattgcacaaa	32400
attcaggttc	ctcacagacc	tgtataactg	tcaaaagcgg	aacattaaca	aaaataccgc	32460
gatcccgtag	gtcccttcgc	agggccagct	gaacataatc	gtgcaggtct	gcacggacca	32520
gcgcggccac	ttccccgcca	ggaaccttga	caaaagaacc	cacactgatt	atgacacgca	32580
tactcggagc	tatgctaacc	agcgtagccc	cgatgtaagc	tttgttgcat	gggcggcgat	32640
ataaaatgca	agggtgctgct	caaaaaatca	ggcaaagcct	cgcgcaaaaa	agaaagcaca	32700
tcgtagtcat	gctcatgcag	ataaaggcag	gtaagctccg	gaaccaccac	agaaaaagac	32760
accatttttc	tctcaaaccat	gtctgcgggt	ttctgcataa	acacaaaata	aaataacaaa	32820
aaaacattta	aacattagaa	gcctgtctta	caacaggaaa	aacaaccctt	ataagcataa	32880
gacggactac	ggccatgccg	gcgtgaccgt	aaaaaaaactg	gtcaccgtga	ttaaaaagca	32940
ccaccgacag	ctcctcggtc	atgtccggag	tcataatgta	agactcggtg	aacacatcag	33000
gttgattcat	cggtcagtg	taaaaagcga	ccgaaatagc	ccgggggaat	acatacccg	33060
aggcgtagag	acaacattac	agcccccata	ggaggtataa	caaaattaat	aggagagaaa	33120
aacacataaa	cacctgaaaa	accctcctgc	ctaggcaaaa	tagcaccctc	ccgctccaga	33180
ccaacataca	gcgcttcaca	gcggcagcct	aacagtcagc	cttaccagta	aaaaagaaaa	33240
cctattaaaa	aaacaccact	cgacacggca	ccagctcaat	cagtcacagt	gtaaaaaagg	33300
gccaaagtga	gagcgagtat	atataggact	aaaaaatgac	gtaacgggta	aagtccacaa	33360
aaaacaccca	gaaaaccgca	cgcgaaacct	cgcccagaaa	cgaaagccaa	aaaaccacaa	33420
acttcctcaa	atcgctcact	ccgttttccc	acgttacgta	acttcccatt	ttaagaaaaa	33480
tactaattccc	aacacataca	agttactccg	ccctaaaacc	tacgtcacc	gccccgttcc	33540
cacgccccgc	gccacgtcac	aaactccacc	ccctcattat	catattggct	tcaatccaaa	33600
ataaggtata	ttattgatga	tg				33622

<210> 45

<211> 1746

<212> DNA

<213> Artificial Sequence

<220>

<223> 5F KO1

<400> 45

-34-

```

atgaagcgcg caagaccgtc tgaagatacc ttcaaccccg tgtatccata tgacacggaa 60
accggctcct caactgtgcc ttttcttact cctccctttg tatcccccaa tgggtttcaa 120
gagagtcccc ctgggggtact ctctttgcgc ctatccgaac ctctagttac ctccaatggc 180
atgcttgccg tcaaaatggg caacggcctc tctctggacg aggccggcaa ccttacctcc 240
caaaatgtaa ccactgtgag cccacctctc aaaaaaacca agtcaaacat aaacctggaa 300
atatctgcac ccctcacagt tacctcagaa gccctaactg tggctgccgc cgcacctcta 360
atggtcgcg gcaacacact caccatgcaa tcacaggccc cgctaaccgt gcacgactcc 420
aaacttagca ttgccacca aggacccctc acagtgtcag aaggaaagct agccctgcaa 480
acatcaggcc ccctcaccac caccgatagc agtaccctta ctatcactgc ctcacccctc 540
ctaactactg ccactggtag cttgggcatt gacttgaaag agccatttta tacacaaaat 600
ggaaaactag gactaaagta cggggctcct ttgcatgtaa cagacgacct aaacactttg 660
accgtagcaa ctggtccagg tgtgactatt aataactctt ccttgcaaac taaagtact 720
ggagccttgg gttttgattc acaaggcaat atgcaactta atgtagcagg aggactaagg 780
attgattctc aaaacagacg ccttatactt gatgttagtt atccgtttga tgctcaaac 840
caactaaatc taagactagg acagggccct ctttttataa actcagccca caacttggat 900
attaactaca acaaaggcct ttacttgttt acagcttcaa acaattccaa aaagcttgag 960
gttaacctaa gcactgcaa ggggttgatg tttgacgcta cagccatagc cattaatgca 1020
ggagatgggc ttgaatttgg ttcacctaat gcaccaaaca caaatccctc caaaacaaaa 1080
attggccatg gcctagaatt tgattcaaac aaggctatgg ttctaaact aggaactggc 1140
cttagttttg acagcacagg tgccattaca gtaggaaaca aaaataatga taagctaact 1200
ttgtggacca caccagctcc agaggctaac tgtagactaa atgcagagaa agatgctaaa 1260
ctcactttgg tcttaacaaa atgtggcagt caaatacttg ctacagtttc agttttggct 1320
gttaaaggca gtttggtctc aatatctgga acagttcaaa gtgctcatct tattataaga 1380
tttgacgaaa atggagtgt actaaacaat tccttcctgg acccagaata ttggaacttt 1440
agaaatggag atcttactga aggcacagcc tatacaaacg ctgttggtat tatgcctaac 1500
ctatcagctt atccaaaatc tcacggtaaa actgccaaaa gtaacattgt cagtcaagtt 1560
tacttaaacg gagacaaaac taaacctgta acactaacca ttacactaaa cggtacacag 1620
gaaacaggag acacaactcc aagtgcatac tctatgtcat tttcatggga ctggtctggc 1680
cacaactaca ttaatgaaat atttgccaca tcctcttaca ctttttcata cattgcccaa 1740
gaataa

```

<210> 46

<211> 581

<212> PRT

<213> Artificial Sequence

<220>

<223> 5F KO1

<400> 46

```

Met Lys Arg Ala Arg Pro Ser Glu Asp Thr Phe Asn Pro Val Tyr Pro
 1           5           10           15
Tyr Asp Thr Glu Thr Gly Pro Pro Thr Val Pro Phe Leu Thr Pro Pro
          20          25          30
Phe Val Ser Pro Asn Gly Phe Gln Glu Ser Pro Pro Gly Val Leu Ser
          35          40          45
Leu Arg Leu Ser Glu Pro Leu Val Thr Ser Asn Gly Met Leu Ala Leu
          50          55          60
Lys Met Gly Asn Gly Leu Ser Leu Asp Glu Ala Gly Asn Leu Thr Ser
65          70          75          80
Gln Asn Val Thr Thr Val Ser Pro Pro Leu Lys Lys Thr Lys Ser Asn
          85          90          95
Ile Asn Leu Glu Ile Ser Ala Pro Leu Thr Val Thr Ser Glu Ala Leu
          100         105         110
Thr Val Ala Ala Ala Ala Pro Leu Met Val Ala Gly Asn Thr Leu Thr
          115         120         125
Met Gln Ser Gln Ala Pro Leu Thr Val His Asp Ser Lys Leu Ser Ile
130         135         140
Ala Thr Gln Gly Pro Leu Thr Val Ser Glu Gly Lys Leu Ala Leu Gln
145         150         155         160
Thr Ser Gly Pro Leu Thr Thr Thr Asp Ser Ser Thr Leu Thr Ile Thr

```

-35-

Ala	Ser	Pro	Pro	165	Leu	Thr	Thr	Ala	Thr	170	Gly	Ser	Leu	Gly	Ile	175	Asp	Leu
Lys	Glu	Pro	Ile	180	Tyr	Thr	Gln	Asn	185	Gly	Lys	Leu	Gly	Leu	190	Lys	Tyr	Gly
Ala	Pro	195	His	Val	Thr	Asp	Asp	200	Leu	Asn	Thr	Leu	Thr	Val	Ala	Thr		
Gly	Pro	Gly	Val	Thr	Ile	215	Asn	Asn	Thr	Ser	Leu	Gln	Thr	Lys	Val	Thr		
225	Gly	Ala	Leu	Gly	Phe	230	Asp	Ser	Gln	Gly	Asn	Met	Gln	Leu	Asn	Val	Ala	
Gly	Gly	Leu	Arg	Ile	245	Asp	Ser	Gln	Asn	Arg	Arg	Leu	Ile	Leu	Asp	Val		
Ser	Tyr	Pro	Phe	260	Asp	Ala	Gln	Asn	Gln	Leu	Asn	Leu	Arg	Leu	Gly	Gln		
Gly	Pro	Leu	Phe	275	Ile	Asn	Ser	Ala	His	Asn	Leu	Asp	Ile	Asn	Tyr	Asn		
Lys	Gly	Leu	Tyr	Leu	Phe	310	Thr	Ala	Ser	Asn	Asn	Ser	Lys	Lys	Leu	Glu		
305	Val	Asn	Leu	Ser	Thr	325	Ala	Lys	Gly	Leu	Met	Phe	Asp	Ala	Thr	Ala	Ile	
Ala	Ile	Asn	Ala	Gly	Asp	Gly	Leu	Glu	Phe	Gly	Ser	Pro	Asn	Ala	Pro			
Asn	Thr	Asn	Pro	Leu	Lys	Thr	Lys	Ile	Gly	His	Gly	Leu	Glu	Phe	Asp			
Ser	Asn	Lys	Ala	Met	Val	Pro	Lys	Leu	Gly	Thr	Gly	Leu	Ser	Phe	Asp			
370	Ser	Thr	Gly	Ala	Ile	Thr	Val	Gly	Asn	Lys	Asn	Asn	Asp	Lys	Leu	Thr		
385	Leu	Trp	Thr	Thr	Pro	405	Ala	Pro	Glu	Ala	Asn	Cys	Arg	Leu	Asn	Ala	Glu	
Lys	Asp	Ala	Lys	Leu	Thr	Leu	Val	Leu	Thr	Lys	Cys	Gly	Ser	Gln	Ile			
Leu	Ala	Thr	Val	Ser	Val	Leu	Ala	Val	Lys	Gly	Ser	Leu	Ala	Pro	Ile			
Ser	Gly	Thr	Val	Gln	Ser	Ala	His	Leu	Ile	Ile	Arg	Phe	Asp	Glu	Asn			
450	Gly	Val	Leu	Leu	Asn	Asn	Ser	Phe	Leu	Asp	Pro	Glu	Tyr	Trp	Asn	Phe		
465	Arg	Asn	Gly	Asp	Leu	Thr	Glu	Gly	Thr	Ala	Tyr	Thr	Asn	Ala	Val	Gly		
Phe	Met	Pro	Asn	Leu	Ser	Ala	Tyr	Pro	Lys	Ser	His	Gly	Lys	Thr	Ala			
Lys	Ser	Asn	Ile	Val	Ser	Gln	Val	Tyr	Leu	Asn	Gly	Asp	Lys	Thr	Lys			
Pro	Val	Thr	Leu	Thr	Ile	Thr	Leu	Asn	Gly	Thr	Gln	Glu	Thr	Gly	Asp			
530	Thr	Thr	Pro	Ser	Ala	Tyr	Ser	Met	Ser	Phe	Ser	Trp	Asp	Trp	Ser	Gly		
545	His	Asn	Tyr	Ile	Asn	Glu	Ile	Phe	Ala	Thr	Ser	Ser	Tyr	Thr	Phe	Ser		
Tyr	Ile	Ala	Gln	Glu														

<210> 47

<211> 1776

<212> DNA

<213> Artificial Sequence

<220>

-36-

<223> 5F KO1RGD

<400> 47

```

atgaagcgcg caagaccgtc tgaagatacc ttcaaccccg tgtatccata tgacacggaa 60
accggtcctc caactgtgcc ttttcttact cctccctttg tatcccccaa tgggtttcaa 120
gagagtcccc ctgggggtact ctcttttgcgc ctatccgaac ctctagttac ctccaatggc 180
atgctttgcgc tcaaaatggg caacggcctc tctctggacg aggcgggcaa ccttacctcc 240
caaaatgtaa ccactgtgag cccacctctc aaaaaaacca agtcaaacat aaacctggaa 300
atatctgcac ccctcacagt tacctcagaa gccctaactg tggctgccgc cgcacctcta 360
atggctcgcgg gcaacacact caccatgcaa tcacaggccc cgctaaccgt gcacgactcc 420
aaacttagca ttgccaccca aggaccctc acagtgtcag aaggaaagct agccctgcaa 480
acatcaggcc ccctcaccac caccgatagc agtaccctta ctatcactgc ctcacccctc 540
ctaactactg ccactggtag cttgggcatt gacttgaaag agcccattta tacacaaaat 600
ggaaaactag gactaaagta cggggctcct ttgcatgtaa cagacgacct aaacactttg 660
accgtagcaa ctggtccagg tgtgactatt aataatactt ccttgcaaac taaagtact 720
ggagccttgg gttttgattc acaaggcaat atgcaactta atgtagcagg aggactaagg 780
attgattctc aaaaacagacg ccttatactt gatgttagtt atccgtttga tgctcaaaac 840
caactaaatc taagactagg acagggcctt ctttttataa actcagccca caacttggat 900
attaactaca acaaaggcct ttacttgttt acagcttcaa acaattccaa aaagcttgag 960
gttaacctaa gcactgcaa ggggttgatg tttgacgcta cagccatagc cattaatgca 1020
ggagatgggc ttgaatttgg ttcacctaat gcaccaaaca caaatcccct caaaacaaaa 1080
attggccatg gcctagaatt tgattcaaac aaggctatgg ttcctaaact aggaactggc 1140
cttagttttg acagcacagg tgccattaca gtaggaaaca aaaataatga taagctaact 1200
ttgtggacca caccagctcc atctoctaac tgtagactaa atgcagagaa agatgctaaa 1260
ctcacttttg tcttaacaaa atgtggcagt caaatacttg ctacagtttc agttttggct 1320
gttaaaggca gtttggtctc aatatctgga acagttcaaa gtgctcatct tattataaga 1380
tttgacgaaa atggagtgtc actaaacaat tccttctctg acccagaata ttggaacttt 1440
agaaatggag atcttactga aggcacagcc tatacaaacg ctgttggtatt tatgcctaac 1500
ctatcagctt atccaaaatc tcacggtaaa actgccaaaa gtaacattgt cagtcaagtt 1560
tacttaaacg gagacaaaac taaacctgta acactaacca ttactactaa cggtagacag 1620
gaaacaggtg atcatttgtg ttgtcgtggt gattgttttt gtacaactcc aagtgcatac 1680
tctatgtcat tttcatggga ctggtctggc cacaactaca ttaatgaaat atttgccaca 1740
tcctcttaca ctttttcata cattgcccac gaataa 1776

```

<210> 48

<211> 591

<212> PRT

<213> Artificial Sequence

<220>

<223> 5F KO1RGD

<400> 48

```

Met Lys Arg Ala Arg Pro Ser Glu Asp Thr Phe Asn Pro Val Tyr Pro
 1           5           10           15
Tyr Asp Thr Glu Thr Gly Pro Pro Thr Val Pro Phe Leu Thr Pro Pro
          20          25          30
Phe Val Ser Pro Asn Gly Phe Gln Glu Ser Pro Pro Gly Val Leu Ser
          35          40          45
Leu Arg Leu Ser Glu Pro Leu Val Thr Ser Asn Gly Met Leu Ala Leu
          50          55          60
Lys Met Gly Asn Gly Leu Ser Leu Asp Glu Ala Gly Asn Leu Thr Ser
65          70          75          80
Gln Asn Val Thr Thr Val Ser Pro Pro Leu Lys Lys Thr Lys Ser Asn
          85          90          95
Ile Asn Leu Glu Ile Ser Ala Pro Leu Thr Val Thr Ser Glu Ala Leu
          100         105         110
Thr Val Ala Ala Ala Ala Pro Leu Met Val Ala Gly Asn Thr Leu Thr
          115         120         125
Met Gln Ser Gln Ala Pro Leu Thr Val His Asp Ser Lys Leu Ser Ile
          130         135         140

```

-37-

Ala	Thr	Gln	Gly	Pro	Leu	Thr	Val	Ser	Glu	Gly	Lys	Leu	Ala	Leu	Gln
145					150					155					160
Thr	Ser	Gly	Pro	Leu	Thr	Thr	Thr	Asp	Ser	Ser	Thr	Leu	Thr	Ile	Thr
				165					170						175
Ala	Ser	Pro	Pro	Leu	Thr	Thr	Ala	Thr	Gly	Ser	Leu	Gly	Ile	Asp	Leu
			180						185					190	
Lys	Glu	Pro	Ile	Tyr	Thr	Gln	Asn	Gly	Lys	Leu	Gly	Leu	Lys	Tyr	Gly
	195						200					205			
Ala	Pro	Leu	His	Val	Thr	Asp	Asp	Leu	Asn	Thr	Leu	Thr	Val	Ala	Thr
	210					215						220			
Gly	Pro	Gly	Val	Thr	Ile	Asn	Asn	Thr	Ser	Leu	Gln	Thr	Lys	Val	Thr
225					230						235				240
Gly	Ala	Leu	Gly	Phe	Asp	Ser	Gln	Gly	Asn	Met	Gln	Leu	Asn	Val	Ala
				245					250					255	
Gly	Gly	Leu	Arg	Ile	Asp	Ser	Gln	Asn	Arg	Arg	Leu	Ile	Leu	Asp	Val
			260						265				270		
Ser	Tyr	Pro	Phe	Asp	Ala	Gln	Asn	Gln	Leu	Asn	Leu	Arg	Leu	Gly	Gln
	275						280					285			
Gly	Pro	Leu	Phe	Ile	Asn	Ser	Ala	His	Asn	Leu	Asp	Ile	Asn	Tyr	Asn
	290						295				300				
Lys	Gly	Leu	Tyr	Leu	Phe	Thr	Ala	Ser	Asn	Asn	Ser	Lys	Lys	Leu	Glu
305					310					315					320
Val	Asn	Leu	Ser	Thr	Ala	Lys	Gly	Leu	Met	Phe	Asp	Ala	Thr	Ala	Ile
				325					330					335	
Ala	Ile	Asn	Ala	Gly	Asp	Gly	Leu	Glu	Phe	Gly	Ser	Pro	Asn	Ala	Pro
			340					345					350		
Asn	Thr	Asn	Pro	Leu	Lys	Thr	Lys	Ile	Gly	His	Gly	Leu	Glu	Phe	Asp
	355						360					365			
Ser	Asn	Lys	Ala	Met	Val	Pro	Lys	Leu	Gly	Thr	Gly	Leu	Ser	Phe	Asp
	370					375					380				
Ser	Thr	Gly	Ala	Ile	Thr	Val	Gly	Asn	Lys	Asn	Asn	Asp	Lys	Leu	Thr
385					390					395					400
Leu	Trp	Thr	Thr	Pro	Ala	Pro	Glu	Ala	Asn	Cys	Arg	Leu	Asn	Ala	Glu
				405					410					415	
Lys	Asp	Ala	Lys	Leu	Thr	Leu	Val	Leu	Thr	Lys	Cys	Gly	Ser	Gln	Ile
			420					425					430		
Leu	Ala	Thr	Val	Ser	Val	Leu	Ala	Val	Lys	Gly	Ser	Leu	Ala	Pro	Ile
	435						440					445			
Ser	Gly	Thr	Val	Gln	Ser	Ala	His	Leu	Ile	Ile	Arg	Phe	Asp	Glu	Asn
	450					455					460				
Gly	Val	Leu	Leu	Asn	Asn	Ser	Phe	Leu	Asp	Pro	Glu	Tyr	Trp	Asn	Phe
465					470					475					480
Arg	Asn	Gly	Asp	Leu	Thr	Glu	Gly	Thr	Ala	Tyr	Thr	Asn	Ala	Val	Gly
				485					490					495	
Phe	Met	Pro	Asn	Leu	Ser	Ala	Tyr	Pro	Lys	Ser	His	Gly	Lys	Thr	Ala
			500					505					510		
Lys	Ser	Asn	Ile	Val	Ser	Gln	Val	Tyr	Leu	Asn	Gly	Asp	Lys	Thr	Lys
		515					520					525			
Pro	Val	Thr	Leu	Thr	Ile	Thr	Leu	Asn	Gly	Thr	Gln	Glu	Thr	Gly	Asp
	530					535						540			
His	Cys	Asp	Cys	Arg	Gly	Asp	Cys	Phe	Cys	Thr	Thr	Pro	Ser	Ala	Tyr
545					550					555					560
Ser	Met	Ser	Phe	Ser	Trp	Asp	Trp	Ser	Gly	His	Asn	Tyr	Ile	Asn	Glu
				565					570					575	
Ile	Phe	Ala	Thr	Ser	Ser	Tyr	Thr	Phe	Ser	Tyr	Ile	Ala	Gln	Glu	
			580					585					590		

<210> 49
 <211> 1746
 <212> DNA

-38-

<213> Artificial Sequence

<220>

<223> 5F KO12

<400> 49

```

atgaagcgcg caagaccgctc tgaagatacc ttcaaccccg tgtatccata tgacacggaa 60
accggtcctc caactgtgcc ttttcttact cctccctttg tatcccccac tgggtttcaa 120
gagagtcccc ctgggggtact ctctttgctc ctatccgaac ctctagttac ctccaatggc 180
atgcttgctc tcaaaatggg caacggcctc tctctggacg aggcgggcaa ccttacctcc 240
caaaatgtaa ccactgtgag cccacctctc aaaaaaacca agtcaaacat aaacctggaa 300
atgctgctac ccctcacagt tacctcagaa gccctaactg tggctgccgc cgcacctcta 360
atggctgcgg gcaacacact caccatgcaa tcacaggccc cgctaaccgt gcacgactcc 420
aaacttagca ttgccaccca aggaccctc acagtgtcag aaggaaagct agccctgcaa 480
acatcaggcc ccctcaccac caccgatagc agtaccctta ctatcactgc ctacccccct 540
ctaactactg ccactggtag cttgggcatg gacttgaaag agcccattta tacacaaaat 600
ggaaaactag gactaaagta cggggctcct ttgcatgtaa cagacgacct aaacactttg 660
accgtagcaa ctgggtccagg tgtgactatt aataatactt ccttgcaaac taaagttagt 720
ggagccttgg gttttgattc acaaggcaat atgcaactta atgtagcagg aggactaagg 780
attgattctc aaaacagacg ccttatactt gatgttagtt atccgtttga tgctcaaaac 840
caactaaatc taagactagg acagggccct ctttttataa actcagcccc caacttggat 900
attaactaca acaaaggcct ttacttgttt acagcttcaa acaattocaa aaagcttgag 960
gttaacctaa gcaactgcaa ggggttgatg tttgacgcta cagccatagc cattaatgca 1020
ggagatgggg ttgaatttgg ttcacctaat gcaccaaaca caaatccctt caaaacaaaa 1080
attggccatg gcctagaatt tgattcaaac aaggctatgg ttcctaaact aggaactggc 1140
cttagttttg acagcacagg tgccattaca gtaggaaaca aaaataatga taagctaact 1200
ttgtggacca caccagctcc atctcctaac tgttactaa atggaggcgg agatgctaaa 1260
ctcacttttg tcttaacaaa atgtggcagt caaatacttg ctacagtttc agttttggct 1320
gttaaaggca gtttggtccc aatatctgga acagtccaac gtgctcatct tattataaga 1380
tttgacgaaa atggagtgtc actaaacaat tccttcctgg acccagaata ttggaacttt 1440
agaaatggag atcttactga aggcacagcc tatacaaacg ctggttgatt tatgcctaac 1500
ctatcagctt atccaaaatc tcacggtaaa actgccaaaa gtaacattgt cagtcaagtt 1560
tacttaaacg gagacaaaac taaacctgta acactaacca ttactactaa cgggtacacag 1620
gaaacaggag acacaactcc aagtgcatac tctatgtcat tttcatggga ctgggtctggc 1680
cacaactaca ttaatgaaat atttgccaca tcctcttaca ctttttcata cattgccccaa 1740
gaataa

```

<210> 50

<211> 581

<212> PRT

<213> Artificial Sequence

<220>

<223> 5F KO12

<400> 50

```

Met Lys Arg Ala Arg Pro Ser Glu Asp Thr Phe Asn Pro Val Tyr Pro
1      5      10      15
Tyr Asp Thr Glu Thr Gly Pro Pro Thr Val Pro Phe Leu Thr Pro Pro
20      25      30
Phe Val Ser Pro Asn Gly Phe Gln Glu Ser Pro Pro Gly Val Leu Ser
35      40      45
Leu Arg Leu Ser Glu Pro Leu Val Thr Ser Asn Gly Met Leu Ala Leu
50      55      60
Lys Met Gly Asn Gly Leu Ser Leu Asp Glu Ala Gly Asn Leu Thr Ser
65      70      75      80
Gln Asn Val Thr Thr Val Ser Pro Pro Leu Lys Lys Thr Lys Ser Asn
85      90      95
Ile Asn Leu Glu Ile Ser Ala Pro Leu Thr Val Thr Ser Glu Ala Leu
100     105     110
Thr Val Ala Ala Ala Ala Pro Leu Met Val Ala Gly Asn Thr Leu Thr

```

[illegible]

-40-

<210> 51
 <211> 1746
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> 5F S*

<400> 51
 atgaagcgcg caagaccgtc tgaagatacc ttcaaccccg tgtatccata tgacacggaa 60
 accggtcctc caactgtgcc ttttcttact cctccctttg tatcccccaa tggggttcaa 120
 gagagtcccc ctgggggtact ctctttgcgc ctatccgaac ctctagttac ctccaatggc 180
 atgcttgcgc tcaaaatggg caacggcctc tctctggacg aggccggcaa ccttacctcc 240
 caaaatgtaa ccactgtgag cccacctctc ggagccggag cctcaaacat aaacctggaa 300
 atatctgcac ccctcacagt tacctcagaa gccctaactg tgggtgccgc cgcacctcta 360
 atggtcgcgg gcaacacact caccatgcaa tcacaggccc cgctaaccgt gcacgactcc 420
 aaacttagca ttgccaccca aggaccctc acagtgtcag aaggaaagct agccctgcaa 480
 acatcaggcc ccctcaccac caccgatagc agtaccctta ctatcactgc ctcaccccct 540
 ctaactactg ccactggtag cttgggcatt gacttgaaag agcccattta tacacaaaat 600
 ggaaaactag gactaaagta cggggctcct ttgcatgtaa cagacgacct aaacactttg 660
 accgtagcaa ctgggtccagg tgtgactatt aataatactt ccttgcaaac taaagttact 720
 ggagccttgg gttttgattc acaaggcaat atgcaactta atgtagcagg aggactaagg 780
 attgattctc aaaacagacg ccttatactt gatgttagtt atccgtttga tgctcaaaac 840
 caactaaatc taagactagg acagggcctt ctttttataa actcagccca caacttggat 900
 attaactaca acaaaggcct ttactttgtt acagcttcaa acaattccaa aaagcttgag 960
 gttaacctaa gcactgcaa ggggttgatg tttgacgcta cagccatagc cattaatgca 1020
 ggagatgggc ttgaatttgg ttcacctaata gcaccaaaca caaatcccct caaaacaaaa 1080
 attggccatg gcctagaatt tgattcaaac aaggctatgg ttcttaaact aggaactggc 1140
 cttagttttg acagcacagg tgccattaca gtaggaaaca aaaataatga taagctaact 1200
 ttgtggacca caccagctcc atctcctaac tgtagactaa atgcagagaa agatgctaaa 1260
 ctcactttgg tcttaacaaa atgtggcagt caaatacttg ctacagtttc agttttgggt 1320
 gttaaaggca gtttgggtcc aatatctgga acagttcaaa gtgctcatct tattataaga 1380
 tttgacgaaa atggagtgtc actaaacaat tccttcctgg acccagaata ttggaacttt 1440
 agaaatggag atcttactga aggcacagcc tatacaaacg ctgttggatt tatgcctaac 1500
 ctatcagctt atccaaaatc tcacggtaaa actgccaaaa gtaacattgt cagtcaagtt 1560
 tacttaaacg gagacaaaac taaacctgta acactaacca ttacactaaa cggtagacag 1620
 gaaacaggag acacaactcc aagtgcatac tctatgtcat tttcatggga ctgggtctggc 1680
 cacaactaca ttaatgaaat atttgccaca tcctcttaca ctttttcata cattgcccac 1740
 gaataa 1746

<210> 52
 <211> 581
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> 5F S*

<400> 52
 Met Lys Arg Ala Arg Pro Ser Glu Asp Thr Phe Asn Pro Val Tyr Pro
 1 5 10 15
 Tyr Asp Thr Glu Thr Gly Pro Pro Thr Val Pro Phe Leu Thr Pro Pro
 20 25 30
 Phe Val Ser Pro Asn Gly Phe Gln Glu Ser Pro Pro Gly Val Leu Ser
 35 40 45
 Leu Arg Leu Ser Glu Pro Leu Val Thr Ser Asn Gly Met Leu Ala Leu
 50 55 60
 Lys Met Gly Asn Gly Leu Ser Leu Asp Glu Ala Gly Asn Leu Thr Ser
 65 70 75 80
 Gln Asn Val Thr Thr Val Ser Pro Pro Leu Gly Ala Gly Ala Ser Asn
 85 90 95

Ile	Asn	Leu	Glu	Ile	Ser	Ala	Pro	Leu	Thr	Val	Thr	Ser	Glu	Ala	Leu
Thr	Val	Ala	Ala	Ala	Ala	Pro	Leu	Met	Val	Ala	Gly	Asn	Thr	Leu	Thr
Met	Gln	Ser	Gln	Ala	Pro	Leu	Thr	Val	His	Asp	Ser	Lys	Leu	Ser	Ile
Ala	Thr	Gln	Gly	Pro	Leu	Thr	Val	Ser	Glu	Gly	Lys	Leu	Ala	Leu	Gln
145				150						155					160
Thr	Ser	Gly	Pro	Leu	Thr	Thr	Thr	Asp	Ser	Ser	Thr	Leu	Thr	Ile	Thr
				165						170					175
Ala	Ser	Pro	Pro	Leu	Thr	Thr	Ala	Thr	Gly	Ser	Leu	Gly	Ile	Asp	Leu
				180						185					190
Lys	Glu	Pro	Ile	Tyr	Thr	Gln	Asn	Gly	Lys	Leu	Gly	Leu	Lys	Tyr	Gly
				195											200
Ala	Pro	Leu	His	Val	Thr	Asp	Asp	Leu	Asn	Thr	Leu	Thr	Val	Ala	Thr
															210
Gly	Pro	Gly	Val	Thr	Ile	Asn	Asn	Thr	Ser	Leu	Gln	Thr	Lys	Val	Thr
225					230										240
Gly	Ala	Leu	Gly	Phe	Asp	Ser	Gln	Gly	Asn	Met	Gln	Leu	Asn	Val	Ala
				245											255
Gly	Gly	Leu	Arg	Ile	Asp	Ser	Gln	Asn	Arg	Arg	Leu	Ile	Leu	Asp	Val
				260											270
Ser	Tyr	Pro	Phe	Asp	Ala	Gln	Asn	Gln	Leu	Asn	Leu	Arg	Leu	Gly	Gln
				275											285
Gly	Pro	Leu	Phe	Ile	Asn	Ser	Ala	His	Asn	Leu	Asp	Ile	Asn	Tyr	Asn
															290
Lys	Gly	Leu	Tyr	Leu	Phe	Thr	Ala	Ser	Asn	Asn	Ser	Lys	Lys	Leu	Glu
305					310										320
Val	Asn	Leu	Ser	Thr	Ala	Lys	Gly	Leu	Met	Phe	Asp	Ala	Thr	Ala	Ile
				325											335
Ala	Ile	Asn	Ala	Gly	Asp	Gly	Leu	Glu	Phe	Gly	Ser	Pro	Asn	Ala	Pro
				340											350
Asn	Thr	Asn	Pro	Leu	Lys	Thr	Lys	Ile	Gly	His	Gly	Leu	Glu	Phe	Asp
				355											365
Ser	Asn	Lys	Ala	Met	Val	Pro	Lys	Leu	Gly	Thr	Gly	Leu	Ser	Phe	Asp
															370
Ser	Thr	Gly	Ala	Ile	Thr	Val	Gly	Asn	Lys	Asn	Asp	Lys	Leu	Thr	400
385					390										
Leu	Trp	Thr	Thr	Pro	Ala	Pro	Ser	Pro	Asn	Cys	Arg	Leu	Asn	Ala	Glu
				405											415
Lys	Asp	Ala	Lys	Leu	Thr	Leu	Val	Leu	Thr	Lys	Cys	Gly	Ser	Gln	Ile
				420											430
Leu	Ala	Thr	Val	Ser	Val	Leu	Ala	Val	Lys	Gly	Ser	Leu	Ala	Pro	Ile
				435											445
Ser	Gly	Thr	Val	Gln	Ser	Ala	His	Leu	Ile	Ile	Arg	Phe	Asp	Glu	Asn
															450
Gly	Val	Leu	Leu	Asn	Asn	Ser	Phe	Leu	Asp	Pro	Glu	Tyr	Trp	Asn	Phe
465															

-42-

580

<210> 53
 <211> 1776
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> 5F S*RGD

<400> 53
 atgaagcgcg caagaccgtc tgaagatacc ttcaaccccg tgtatccata tgacacggaa 60
 accggtcctc caactgtgcc ttttcttact cctccctttg tatcccccaa tgggtttcaa 120
 gagagtcccc ctgggggtact ctctttgcgc ctatccgaac ctctagttac ctccaatggc 180
 atgcttgcgc tcaaaatggg caacggcctc tctctggacg aggccggcaa ccttacctcc 240
 caaaatgtaa ccactgtgag cccacctctc ggagccggag cctcaaacad aaacctggaa 300
 atatctgcac ccctcacagt tacctcagaa gccctaactg tggctgcccgc cgcacctcta 360
 atggctgcgc gcaacacact caccatgcaa tcacaggccc cgctaaccgt gcacgactcc 420
 aaacttagca ttgccaccca aggaccctc acagtgtcag aaggaaagct agccctgcaa 480
 acatcaggcc ccctcaccac caccgatagc agtaccctta ctatcactgc ctcacccctc 540
 ctaactactg ccactggtag cttgggcatt gacttgaaag agcccattta tacacaaaat 600
 ggaaaactag gactaaagta cggggctcct ttgcatgtaa cagacgacct aaacactttg 660
 accgtagcaa ctgggtccagg tgtgactatt aataatactt ccttgcaaac taaagtact 720
 ggagccttgg gttttgattc acaaggcaat atgcaactta atgtagcagg aggactaagg 780
 attgattctc aaaacagacg ccttataact gatgttagtt atccgtttga tgctcaaac 840
 caactaaatc taagactagg acagggcctt ctttttataa actcagccca caacttggat 900
 attaactaca acaaaggcct ttacttggtt acagcttcaa acaattccaa aaagcttgag 960
 gttaacctaa gcaactgcaa ggggttgatg tttgacgcta cagccatagc cattaatgca 1020
 ggagatgggc ttgaatttgg ttcacctaata gcaccaaaca caaatcccct caaaacaaaa 1080
 attggccatg gcctagaatt tgattcaaac aaggctatgg ttccctaaact aggaactggc 1140
 cttagttttg acagcacagg tgccattaca gtaggaaaca aaaataatga taagctaact 1200
 ttgtggacca caccagctcc atctcctaac tgtagactaa atgcagagaa agatgctaaa 1260
 ctcacttttg tcttaacaaa atgtggcagt caaatacttg ctacagtttc agttttggct 1320
 gttaaaggca gtttggctcc aatatctgga acagttcaaa gtgctcatct tattataaga 1380
 tttgacgaaa atggagtgt actaaacaat tccttcctgg acccagaata ttggaacttt 1440
 agaaatggag atcttactga aggcacagcc tatacaaacg ctggttgatt tatgcctaac 1500
 ctatcagctt atccaaaatc tcacggtaaa actgcaaaaa gtaacattgt cagtcaagtt 1560
 tacttaaacg gagacaaaaa taaacctgta acactaacca ttactactaa cggtacacag 1620
 gaaacaggtg atcattgtga ttgtcgtggg gattgttttt gtacaactcc aagtgcatac 1680
 tctatgtcat tttcatggga ctggctctggc cacaactaca ttaatgaaat atttgccaca 1740
 tcctcttaca ctttttcata cattgcccac gaataa 1776

<210> 54
 <211> 591
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> 5F S*RGD

<400> 54
 Met Lys Arg Ala Arg Pro Ser Glu Asp Thr Phe Asn Pro Val Tyr Pro
 1 5 10 15
 Tyr Asp Thr Glu Thr Gly Pro Pro Thr Val Pro Phe Leu Thr Pro Pro
 20 25 30
 Phe Val Ser Pro Asn Gly Phe Gln Glu Ser Pro Pro Gly Val Leu Ser
 35 40 45
 Leu Arg Leu Ser Glu Pro Leu Val Thr Ser Asn Gly Met Leu Ala Leu
 50 55 60
 Lys Met Gly Asn Gly Leu Ser Leu Asp Glu Ala Gly Asn Leu Thr Ser

-43-

65	Gln	Asn	Val	Thr	Thr	70	Val	Ser	Pro	Pro	Leu	75	Gly	Ala	Gly	Ala	Ser	80	Asn
					85						90						95		
Ile	Asn	Leu	Glu	Ile	Ser	Ala	Pro	Leu	Thr	Val	Thr	Val	Thr	Ser	Glu	Ala	Leu		
			100					105							110				
Thr	Val	Ala	Ala	Ala	Pro	Leu	Met	Val	Ala	Gly	Asn	Thr	Leu	Thr					
			115					120						125					
Met	Gln	Ser	Gln	Ala	Pro	Leu	Thr	Val	His	Asp	Ser	Lys	Leu	Ser	Ile				
			130					135					140						
Ala	Thr	Gln	Gly	Pro	Leu	Thr	Val	Ser	Glu	Gly	Lys	Leu	Ala	Leu	Gln				
145								150					155						
Thr	Ser	Gly	Pro	Leu	Thr	Thr	Thr	Asp	Ser	Ser	Thr	Leu	Thr	Ile	Thr				
								165						170					
Ala	Ser	Pro	Pro	Leu	Thr	Thr	Ala	Thr	Gly	Ser	Leu	Gly	Ile	Asp	Leu				
			180					185						190					
Lys	Glu	Pro	Ile	Tyr	Thr	Gln	Asn	Gly	Lys	Leu	Gly	Leu	Lys	Tyr	Gly				
			195					200					205						
Ala	Pro	Leu	His	Val	Thr	Asp	Asp	Leu	Asn	Thr	Leu	Thr	Val	Ala	Thr				
			210					215					220						
Gly	Pro	Gly	Val	Thr	Ile	Asn	Asn	Thr	Ser	Leu	Gln	Thr	Lys	Val	Thr				
225								230					235						
Gly	Ala	Leu	Gly	Phe	Asp	Ser	Gln	Gly	Asn	Met	Gln	Leu	Asn	Val	Ala				
								245						250					
Gly	Gly	Leu	Arg	Ile	Asp	Ser	Gln	Asn	Arg	Arg	Leu	Ile	Leu	Asp	Val				
			260					265						270					
Ser	Tyr	Pro	Phe	Asp	Ala	Gln	Asn	Gln	Leu	Asn	Leu	Arg	Leu	Gly	Gln				
			275					280					285						
Gly	Pro	Leu	Phe	Ile	Asn	Ser	Ala	His	Asn	Leu	Asp	Ile	Asn	Tyr	Asn				
			290					295					300						
Lys	Gly	Leu	Tyr	Leu	Phe	Thr	Ala	Ser	Asn	Asn	Ser	Lys	Lys	Leu	Glu				
305								310					315						
Val	Asn	Leu	Ser	Thr	Ala	Lys	Gly	Leu	Met	Phe	Asp	Ala	Thr	Ala	Ile				
								325						330					
Ala	Ile	Asn	Ala	Gly	Asp	Gly	Leu	Glu	Phe	Gly	Ser	Pro	Asn	Ala	Pro				
			340					345						350					
Asn	Thr	Asn	Pro	Leu	Lys	Thr	Lys	Ile	Gly	His	Gly	Leu	Glu	Phe	Asp				
			355					360					365						
Ser	Asn	Lys	Ala	Met	Val	Pro	Lys	Leu	Gly	Thr	Gly	Leu	Ser	Phe	Asp				
			370					375					380						
Ser	Thr	Gly	Ala	Ile	Thr	Val	Gly	Asn	Lys	Asn	Asn	Asp	Lys	Leu	Thr				
385								390					395						
Leu	Trp	Thr	Thr	Pro	Ala	Pro	Ser	Pro	Asn	Cys	Arg	Leu	Asn	Ala	Glu				
								405						410					
Lys	Asp	Ala	Lys	Leu	Thr	Leu	Val	Leu	Thr	Lys	Cys	Gly	Ser	Gln	Ile				
			420					425						430					
Leu	Ala	Thr	Val	Ser	Val	Leu	Ala	Val	Lys	Gly	Ser	Leu	Ala	Pro	Ile				
			435					440					445						
Ser	Gly	Thr	Val	Gln	Ser	Ala	His	Leu	Ile	Ile	Arg	Phe	Asp	Glu	Asn				
			450					455					460						
Gly	Val	Leu	Leu	Asn	Asn	Ser	Phe	Leu	Asp	Pro	Glu	Tyr	Trp	Asn	Phe				
465								470					475						
Arg	Asn	Gly	Asp	Leu	Thr	Glu	Gly	Thr	Ala	Tyr	Thr	Asn	Ala	Val	Gly				
								485						490					
Phe	Met	Pro	Asn	Leu	Ser	Ala	Tyr	Pro	Lys	Ser	His	Gly	Lys	Thr	Ala				
			500					505					510						
Lys	Ser	Asn	Ile	Val	Ser	Gln	Val	Tyr	Leu	Asn	Gly	Asp	Lys	Thr	Lys				
			515					520					525						
Pro	Val	Thr	Leu	Thr	Ile	Thr	Leu	Asn	Gly	Thr	Gln	Glu	Thr	Gly	Asp				
			530					535					540						
His	Cys	Asp	Cys	Arg	Gly	Asp	Cys	Phe	Cys	Thr	Thr	Pro	Ser	Ala	Tyr				
545								550					555						

-44-

Ser Met Ser Phe Ser Trp Asp Trp Ser Gly His Asn Tyr Ile Asn Glu
 565 570 575
 Ile Phe Ala Thr Ser Ser Tyr Thr Phe Ser Tyr Ile Ala Gln Glu
 580 585 590

<210> 55
 <211> 1746
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> 5F KO1S*

<400> 55
 atgaagcgcg caagaccgtc tgaagatacc ttcaacccccg tgtatccata tgacacggaa 60
 accggctctc caactgtgcc ttttcttact cctccctttg tatcccccaa tgggtttcaa 120
 gagagtcctc ctgggggtact ctctttgctc ctatccgaac ctctagttac ctccaatggc 180
 atgcttgctc tcaaaatggg caacggcctc tctctggacg aggcgggcaa ccttacctcc 240
 caaaatgtaa ccactgtgag cccacctctc ggagccggag cctcaaacat aaacctggaa 300
 atatctgcac ccctcacagt tacctcagaa gccctaactg tggctgcccgc cgcacctcta 360
 atggtcgctg gcaacacact caccatgcaa tcacaggccc cgctaaccgt gcacgactcc 420
 aaacttagca ttggcacccta aggacccctc acagtgtcag aaggaaagct agccctgcaa 480
 acatcaggcc ccctcaccac caccgatagc agtaccctta ctatcactgc ctcaccccct 540
 ctaactactg ccactggtag cttggggcatt gacttgaaaag agcccattta tacacaaaat 600
 ggaaaactag gactaaagta cggggctcct ttgcatgtaa cagacgacct aaacactttg 660
 accgtagcaa ctgggtccagg tgtgactatt aataatactt ccttgcaaac taaagttact 720
 ggagccttgg gttttgattc acaaggcaat atgcaactta atgtagcagg aggactaagg 780
 attgattctc aaaacagacg ccttatactt gatgttagtt atccgtttga tgctcaaaac 840
 caactaaatc taagactagg acagggccct ctttttataa actcagccca caacttggat 900
 attaactaca acaaaggcct ttacttgttt acagcttcaa acaattccaa aaagcttgag 960
 gttaacctaa gcactgccaa ggggttgatg tttgacgcta cagccatagc cattaatgca 1020
 ggagatgggc ttgaatttgg ttcacctaata gcaccaaaca caaatccctt caaaacaaaa 1080
 attggccatg gcctagaatt tgattcaaac aaggctatgg ttcctaaact aggaactggc 1140
 cttagttttg acagcacagg tgccattaca gtaggaaaca aaaataatga taagctaact 1200
 ttgtggacca caccagctcc agaggctaac tgtagactaa atgcagagaa agatgctaaa 1260
 ctacactttg tcttaacaaa atgtggcagt caaatacttg ctacagtttc agttttggct 1320
 gttaaaggca gtttggtctc aatatctgga acagttcaaa gtgctcatct tattataaga 1380
 tttgacgaaa atggagtgct actaaacaat tccttcctgg acccagaata ttggaacttt 1440
 agaaatggag atcttactga aggcacagcc tatacaaacg ctggttgatt tatgcctaac 1500
 ctatcagctt atccaaaatc tcacggtaaa actgccaaaa gtaacattgt cagtcaagtt 1560
 tacttaaacg gagacaaaac taaacctgta acactaacca ttactactaa cgggtacacg 1620
 gaaacaggag acacaactcc aagtgcatac tctatgtcat tttcatggga ctggtctggc 1680
 cacaactaca ttaatgaaat atttgccaca tcctcttaca ctttttcata cattgcccga 1740
 gaataa 1746

<210> 56
 <211> 581
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> 5F KO1S*

<400> 56
 Met Lys Arg Ala Arg Pro Ser Glu Asp Thr Phe Asn Pro Val Tyr Pro
 1 5 10 15
 Tyr Asp Thr Glu Thr Gly Pro Pro Thr Val Pro Phe Leu Thr Pro Pro
 20 25 30
 Phe Val Ser Pro Asn Gly Phe Gln Glu Ser Pro Pro Gly Val Leu Ser
 35 40 45

-45-

Leu Arg Leu Ser Glu Pro Leu Val Thr Ser Asn Gly Met Leu Ala Leu
 50 55 60
 Lys Met Gly Asn Gly Leu Ser Leu Asp Glu Ala Gly Asn Leu Thr Ser
 65 70 75 80
 Gln Asn Val Thr Thr Val Ser Pro Pro Leu Gly Ala Gly Ala Ser Asn
 85 90 95
 Ile Asn Leu Glu Ile Ser Ala Pro Leu Thr Val Thr Ser Glu Ala Leu
 100 105 110
 Thr Val Ala Ala Ala Ala Pro Leu Met Val Ala Gly Asn Thr Leu Thr
 115 120 125
 Met Gln Ser Gln Ala Pro Leu Thr Val His Asp Ser Lys Leu Ser Ile
 130 135 140
 Ala Thr Gln Gly Pro Leu Thr Val Ser Glu Gly Lys Leu Ala Leu Gln
 145 150 155 160
 Thr Ser Gly Pro Leu Thr Thr Thr Asp Ser Ser Thr Leu Thr Ile Thr
 165 170 175
 Ala Ser Pro Pro Leu Thr Thr Ala Thr Gly Ser Leu Gly Ile Asp Leu
 180 185 190
 Lys Glu Pro Ile Tyr Thr Gln Asn Gly Lys Leu Gly Leu Lys Tyr Gly
 195 200 205
 Ala Pro Leu His Val Thr Asp Asp Leu Asn Thr Leu Thr Val Ala Thr
 210 215 220
 Gly Pro Gly Val Thr Ile Asn Asn Thr Ser Leu Gln Thr Lys Val Thr
 225 230 235 240
 Gly Ala Leu Gly Phe Asp Ser Gln Gly Asn Met Gln Leu Asn Val Ala
 245 250 255
 Gly Gly Leu Arg Ile Asp Ser Gln Asn Arg Arg Leu Ile Leu Asp Val
 260 265 270
 Ser Tyr Pro Phe Asp Ala Gln Asn Gln Leu Asn Leu Arg Leu Gly Gln
 275 280 285
 Gly Pro Leu Phe Ile Asn Ser Ala His Asn Leu Asp Ile Asn Tyr Asn
 290 295 300
 Lys Gly Leu Tyr Leu Phe Thr Ala Ser Asn Asn Ser Lys Lys Leu Glu
 305 310 315 320
 Val Asn Leu Ser Thr Ala Lys Gly Leu Met Phe Asp Ala Thr Ala Ile
 325 330 335
 Ala Ile Asn Ala Gly Asp Gly Leu Glu Phe Gly Ser Pro Asn Ala Pro
 340 345 350
 Asn Thr Asn Pro Leu Lys Thr Lys Ile Gly His Gly Leu Glu Phe Asp
 355 360 365
 Ser Asn Lys Ala Met Val Pro Lys Leu Gly Thr Gly Leu Ser Phe Asp
 370 375 380
 Ser Thr Gly Ala Ile Thr Val Gly Asn Lys Asn Asn Asp Lys Leu Thr
 385 390 395 400
 Leu Trp Thr Thr Pro Ala Pro Glu Ala Asn Cys Arg Leu Asn Ala Glu
 405 410 415
 Lys Asp Ala Lys Leu Thr Leu Val Leu Thr Lys Cys Gly Ser Gln Ile
 420 425 430
 Leu Ala Thr Val Ser Val Leu Ala Val Lys Gly Ser Leu Ala Pro Ile
 435 440 445
 Ser Gly Thr Val Gln Ser Ala His Leu Ile Ile Arg Phe Asp Glu Asn
 450 455 460
 Gly Val Leu Leu Asn Asn Ser Phe Leu Asp Pro Glu Tyr Trp Asn Phe
 465 470 475 480
 Arg Asn Gly Asp Leu Thr Glu Gly Thr Ala Tyr Thr Asn Ala Val Gly
 485 490 495
 Phe Met Pro Asn Leu Ser Ala Tyr Pro Lys Ser His Gly Lys Thr Ala
 500 505 510
 Lys Ser Asn Ile Val Ser Gln Val Tyr Leu Asn Gly Asp Lys Thr Lys
 515 520 525
 Pro Val Thr Leu Thr Ile Thr Leu Asn Gly Thr Gln Glu Thr Gly Asp

-46-

530 535 540
 Thr Thr Pro Ser Ala Tyr Ser Met Ser Phe Ser Trp Asp Trp Ser Gly
 545 550 555 560
 His Asn Tyr Ile Asn Glu Ile Phe Ala Thr Ser Ser Tyr Thr Phe Ser
 565 570 575
 Tyr Ile Ala Gln Glu
 580

<210> 57
 <211> 1776
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> 5F KO1S*RGD

<400> 57
 atgaagcgcg caagaccgctc tgaagatacc ttcaaccccg tgtatccata tgacacggaa 60
 accggtcctc caactgtgcc ttttcttact cctccctttg tatcccccaa tggggtttcaa 120
 gagagtcctc ctgggggtact ctcttttgcgc ctatccgaac ctctagttac ctccaatggc 180
 atgcttgcgc tcaaaatggg caacggcctc tctctggacg aggcgggcaa ccttacctcc 240
 caaaatgtaa ccaactgtgag cccacctctc ggagccggag cctcaaacat aaacctggaa 300
 atatctgcac ccctcacagt tacctcagaa gccctaactg tggctgccgc cgcacctcta 360
 atggctcgcg gcaacacact caccatgcaa tcacaggccc cgctaaccgt gcacgactcc 420
 aaacttagca ttgccacca aggaccctc acagtgtcag aaggaaagct agccctgcaa 480
 acatcaggcc ccctcaccac caccgatagc agtaccctta ctatcactgc ctacccccct 540
 ctaactactg ccactggtag cttggggcatt gacttgaaag agcccattta tacacaaaat 600
 ggaaaactag gactaaagta cgggggtcct ttgcatgtaa cagacgacct aaacactttg 660
 accgtagcaa ctggtccagg tgtgactatt aataatactt ccttgcaaac taaagt tact 720
 ggagccttgg gttttgattc acaaggcaat atgcaactta atgtagcagg aggactaagg 780
 attgattctc aaaacagacg ccttatactt gatgttagtt atccgtttga tgctcaaaac 840
 caactaaatc taagactagg acagggcctt ctttttataa actcagccca caacttggat 900
 attaaactaca acaaaggcct ttacttgttt acagcttcaa acaattccaa aaagcttgag 960
 gttaacctaa gcaactgccaa ggggttgatg tttgacgcta cagccatagc cattaatgca 1020
 ggagatgggc ttgaatttgg ttcacctaatt gcaccaaaca caaatccct caaaacaaaa 1080
 attggccatg gcctagaatt tgattcaaac aaggctatgg ttcttaaact aggaactggc 1140
 cttagttttg acagcacagg tgccattaca gtaggaaaca aaaataatga taagctaact 1200
 ttgtggacca caccagctcc agaggctaac tgtagactaa atgcagagaa agatgctaaa 1260
 ctacatttgg tcttaacaaa atgtggcagt caaatacttg ctacagtttc agttttggct 1320
 gttaaaggca gtttggtccc aatatctgga acagttcaaa gtgtcatct tattataaga 1380
 tttgacgaaa atggagtgt actaaacaat tccttcctgg acccagaata ttggaacttt 1440
 agaaatggag atcttactga aggcacagcc tatacaaacg ctgttggatt tatgcctaac 1500
 ctatcagctt atccaaaatc tcacggtaaa actgccaaaa gtaacattgt cagtcaagtt 1560
 tacttaaacg gagacaaaac taaacctgta aactaacca ttacactaaa cggtacacag 1620
 gaaacaggtg atcattgtga ttgtcgtggt gattgttttt gtacaacccc aagtgcatac 1680
 tctatgtcat tttcatggga ctggtctggc cacaactaca ttaatgaaat atttgccaca 1740
 tcctcttaca ctttttcata cattgcccac gaataa 1776

<210> 58
 <211> 591
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> 5F KO1S*RGD

<400> 58
 Met Lys Arg Ala Arg Pro Ser Glu Asp Thr Phe Asn Pro Val Tyr Pro
 1 5 10 15
 Tyr Asp Thr Glu Thr Gly Pro Pro Thr Val Pro Phe Leu Thr Pro Pro

-47-

Phe	Val	Ser	20	Asn	Gly	Phe	Gln	25	Glu	Ser	Pro	Pro	Gly	30	Val	Leu	Ser
		35					40						45				
Leu	Arg	Leu	Ser	Glu	Pro	Leu	Val	Thr	Ser	Asn	Gly	Met	Leu	Ala	Leu		
	50					55					60						
Lys	Met	Gly	Asn	Gly	Leu	Ser	Leu	Asp	Glu	Ala	Gly	Asn	Leu	Thr	Ser		
65					70				75						80		
Gln	Asn	Val	Thr	Thr	Val	Ser	Pro	Pro	Leu	Gly	Ala	Gly	Ala	Ser	Asn		
				85					90					95			
Ile	Asn	Leu	Glu	Ile	Ser	Ala	Pro	Leu	Thr	Val	Thr	Ser	Glu	Ala	Leu		
			100					105					110				
Thr	Val	Ala	Ala	Ala	Ala	Pro	Leu	Met	Val	Ala	Gly	Asn	Thr	Leu	Thr		
		115					120					125					
Met	Gln	Ser	Gln	Ala	Pro	Leu	Thr	Val	His	Asp	Ser	Lys	Leu	Ser	Ile		
	130					135					140						
Ala	Thr	Gln	Gly	Pro	Leu	Thr	Val	Ser	Glu	Gly	Lys	Leu	Ala	Leu	Gln		
145					150				155						160		
Thr	Ser	Gly	Pro	Leu	Thr	Thr	Thr	Asp	Ser	Ser	Thr	Leu	Thr	Ile	Thr		
				165				170						175			
Ala	Ser	Pro	Pro	Leu	Thr	Thr	Ala	Thr	Gly	Ser	Leu	Gly	Ile	Asp	Leu		
			180					185					190				
Lys	Glu	Pro	Ile	Tyr	Thr	Gln	Asn	Gly	Lys	Leu	Gly	Leu	Lys	Tyr	Gly		
	195						200					205					
Ala	Pro	Leu	His	Val	Thr	Asp	Asp	Leu	Asn	Thr	Leu	Thr	Val	Ala	Thr		
	210					215					220						
Gly	Pro	Gly	Val	Thr	Ile	Asn	Asn	Thr	Ser	Leu	Gln	Thr	Lys	Val	Thr		
225					230					235					240		
Gly	Ala	Leu	Gly	Phe	Asp	Ser	Gln	Gly	Asn	Met	Gln	Leu	Asn	Val	Ala		
				245				250						255			
Gly	Gly	Leu	Arg	Ile	Asp	Ser	Gln	Asn	Arg	Arg	Leu	Ile	Leu	Asp	Val		
		260						265				270					
Ser	Tyr	Pro	Phe	Asp	Ala	Gln	Asn	Gln	Leu	Asn	Leu	Arg	Leu	Gly	Gln		
	275					280					285						
Gly	Pro	Leu	Phe	Ile	Asn	Ser	Ala	His	Asn	Leu	Asp	Ile	Asn	Tyr	Asn		
	290				295						300						
Lys	Gly	Leu	Tyr	Leu	Phe	Thr	Ala	Ser	Asn	Asn	Ser	Lys	Lys	Leu	Glu		
305					310				315						320		
Val	Asn	Leu	Ser	Thr	Ala	Lys	Gly	Leu	Met	Phe	Asp	Ala	Thr	Ala	Ile		
				325				330						335			
Ala	Ile	Asn	Ala	Gly	Asp	Gly	Leu	Glu	Phe	Gly	Ser	Pro	Asn	Ala	Pro		
		340					345						350				
Asn	Thr	Asn	Pro	Leu	Lys	Thr	Lys	Ile	Gly	His	Gly	Leu	Glu	Phe	Asp		
		355				360						365					
Ser	Asn	Lys	Ala	Met	Val	Pro	Lys	Leu	Gly	Thr	Gly	Leu	Ser	Phe	Asp		
	370					375					380						
Ser	Thr	Gly	Ala	Ile	Thr	Val	Gly	Asn	Lys	Asn	Asn	Asp	Lys	Leu	Thr		
385					390				395						400		
Leu	Trp	Thr	Thr	Pro	Ala	Pro	Glu	Ala	Asn	Cys	Arg	Leu	Asn	Ala	Glu		
				405				410						415			
Lys	Asp	Ala	Lys	Leu	Thr	Leu	Val	Leu	Thr	Lys	Cys	Gly	Ser	Gln	Ile		
		420						425					430				
Leu	Ala	Thr	Val	Ser	Val	Leu	Ala	Val	Lys	Gly	Ser	Leu	Ala	Pro	Ile		
	435					440						445					
Ser	Gly	Thr	Val	Gln	Ser	Ala	His	Leu	Ile	Ile	Arg	Phe	Asp	Glu	Asn		
	450				455						460						
Gly	Val	Leu	Leu	Asn	Asn	Ser	Phe	Leu	Asp	Pro	Glu	Tyr	Trp	Asn	Phe		
465				470					475					480			
Arg	Asn	Gly	Asp	Leu	Thr	Glu	Gly	Thr	Ala	Tyr	Thr	Asn	Ala	Val	Gly		
				485				490						495			
Phe	Met	Pro	Asn	Leu	Ser	Ala	Tyr	Pro	Lys	Ser	His	Gly	Lys	Thr	Ala		
			500					505					510				

-48-

Lys	Ser	Asn	Ile	Val	Ser	Gln	Val	Tyr	Leu	Asn	Gly	Asp	Lys	Thr	Lys
		515					520					525			
Pro	Val	Thr	Leu	Thr	Ile	Thr	Leu	Asn	Gly	Thr	Gln	Glu	Thr	Gly	Asp
		530				535					540				
His	Cys	Asp	Cys	Arg	Gly	Asp	Cys	Phe	Cys	Thr	Thr	Pro	Ser	Ala	Tyr
545					550				555						560
Ser	Met	Ser	Phe	Ser	Trp	Asp	Trp	Ser	Gly	His	Asn	Tyr	Ile	Asn	Glu
				565					570					575	
Ile	Phe	Ala	Thr	Ser	Ser	Tyr	Thr	Phe	Ser	Tyr	Ile	Ala	Gln	Glu	
			580					585					590		

<210> 59
 <211> 972
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> 35F

<400> 59
 atgaccaaga gagtccggct cagtgactcc ttcaaccctg tctacccta tgaagatgaa 60
 agcacctccc aacacccctt tataaaccca gggtttattt ccccaaattg cttcacacaa 120
 agcccagacg gagttcttac tttaaaatgt ttaacccac taacaaccac aggcggatct 180
 ctacagctaa aagtgggagg gggacttaca gtggatgaca ctgatggtag cttacaagaa 240
 aacatacgtg ctacagcacc cattactaaa aataatcact ctgtagaact atccattgga 300
 aatggattag aaactcaaaa caataaacta tgtgccaaat tgggaaatgg gttaaaattt 360
 aacaacggtg acatttgtat aaaggatagt attaacacct tatggactgg aataaacctt 420
 ccacctaact gtcaaattgt ggaaaacact aatacaaatg atggcaaact tacttttagta 480
 ttagtaaaaa atggagggct tggttaatggc tacgtgtctc tagttgggtg atcagacact 540
 gtgaacccaaa tggtcacaca aaagacagca aacatccaat taagattata ttttgactct 600
 tctggaaatc tattaactga ggaatcagac ttaaaaattc cacttaaaaa taaatcttct 660
 acagcgacca gtgaaactgt agccagcagc aaagccttta tgccaagtac tacagcttat 720
 cccttcaaca ccactactag ggatagttaa aactacattc atggaatatg ttactacatg 780
 actagttatg atagaagtct atttcccttg aacattttcta taatgctaaa cagccgatg 840
 atttcttcca atgttgcccta tgccatacaa tttgaatgga atctaaatgc aagtgaatct 900
 ccagaaagca acatagctac gctgaccaca tccccctttt tcttttctta cattacagaa 960
 gacgacgaat aa 972

<210> 60
 <211> 323
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> 35F

<400> 60
 Met Thr Lys Arg Val Arg Leu Ser Asp Ser Phe Asn Pro Val Tyr Pro
 1 5 10 15
 Tyr Glu Asp Glu Ser Thr Ser Gln His Pro Phe Ile Asn Pro Gly Phe
 20 25 30
 Ile Ser Pro Asn Gly Phe Thr Gln Ser Pro Asp Gly Val Leu Thr Leu
 35 40 45
 Lys Cys Leu Thr Pro Leu Thr Thr Thr Gly Gly Ser Leu Gln Leu Lys
 50 55 60
 Val Gly Gly Gly Leu Thr Val Asp Asp Thr Asp Gly Thr Leu Gln Glu
 65 70 75 80
 Asn Ile Arg Ala Thr Ala Pro Ile Thr Lys Asn Asn His Ser Val Glu
 85 90 95
 Leu Ser Ile Gly Asn Gly Leu Glu Thr Gln Asn Asn Lys Leu Cys Ala

-49-

Lys	Leu	Gly	100	Asn	Gly	Leu	Lys	Phe	105	Asn	Asn	Gly	Asp	110	Ile	Cys	Ile	Lys
Asp	Ser	Ile	115	Asn	Thr	Leu	Trp	Thr	120	Gly	Ile	Asn	Pro	125	Pro	Pro	Asn	Cys
Gln	Ile	Val	130	Glu	Asn	Thr	Asn	Thr	135	Asn	Asp	Gly	Lys	140	Leu	Thr	Leu	Val
145	Leu	Val	Lys	Asn	Gly	Gly	Leu	Val	150	Asn	Gly	Tyr	Val	155	Ser	Leu	Val	Gly
Val	Ser	Asp	Thr	165	Val	Asn	Gln	Met	170	Phe	Thr	Gln	Lys	175	Thr	Ala	Asn	Ile
Gln	Leu	Arg	Leu	180	Tyr	Phe	Asp	Ser	185	Ser	Gly	Asn	Leu	190	Leu	Thr	Glu	Glu
Ser	Asp	Leu	Lys	195	Ile	Pro	Leu	Lys	200	Asn	Lys	Ser	Ser	205	Thr	Ala	Thr	Ser
Glu	Thr	Val	Ala	210	Ser	Ser	Lys	Ala	215	Phe	Met	Pro	Ser	220	Thr	Thr	Ala	Tyr
225	Pro	Phe	Asn	Thr	230	Thr	Arg	Asp	235	Ser	Glu	Asn	Tyr	240	Ile	His	Gly	Ile
Cys	Tyr	Tyr	Met	245	Thr	Ser	Tyr	Asp	250	Arg	Ser	Leu	Phe	255	Pro	Leu	Asn	Ile
Ser	Ile	Met	Leu	260	Asn	Ser	Arg	Met	265	Ile	Ser	Ser	Asn	270	Val	Ala	Tyr	Ala
Ile	Gln	Phe	Glu	275	Trp	Asn	Leu	Asn	280	Ala	Ser	Glu	Ser	285	Pro	Glu	Ser	Asn
Ile	Ala	Thr	Leu	290	Thr	Thr	Ser	Pro	295	Phe	Phe	Phe	Ser	300	Tyr	Ile	Thr	Glu
305	Asp	Asp	Glu	310					315					320				

<210> 61

<211> 1002

<212> DNA

<213> Artificial Sequence

<220>

<223> 35F RGD

<400> 61

atgaccaaga	gagtcgggct	cagtgactcc	ttcaaccctg	tctacccta	tgaagatgaa	60
agcacctccc	aacaccctt	tataaaccca	gggtttat	ccccaaatg	cttcacacaa	120
agccagacg	gagttcttac	tttaaaatgt	ttaacccac	taacaaccac	aggcggtat	180
ctacagctaa	aagtgggagg	gggacttaca	gtggatgaca	ctgatggtag	cttacaagaa	240
aacatacgtg	ctacagcacc	cattactaaa	aataatcact	ctgtagaact	atccattgga	300
aatggattag	aaactcaaaa	caataaacta	tgtgccaaat	tgggaaatg	gttaaaattt	360
aacaacgggtg	acatttgtat	aaaggatagt	attaacacct	tatggactgg	aataaacctt	420
ccacctaact	gtcaaattgt	ggaaaacact	aatacaaatg	atggcaaact	tacttttagta	480
ttagtaaaaa	atggagggtg	tggttaattg	tacgtgtctc	tagttggtgt	atcagacact	540
gtgaacaaaa	tggttcacaca	aaagacagca	aacatccaat	taagattata	ttttgactct	600
tctggaaatc	tattaactga	ggaatcagac	ttaaaaattc	cacttaaaaa	taaatcttct	660
acagcgacca	gtgaaactgt	agccagcagc	aaagccttta	tgccaagtac	tacagcttat	720
cccttcaaca	ccactactag	ggatagtga	aactacattc	atggaatatg	ttactacatg	780
actagttatg	atagaagtct	atttccttg	aacatttcta	taatgctaaa	cagccgtatg	840
atttccttcca	atgtacattg	tgattgtcgt	ggtgattggt	tttgcgcata	tgccatacaa	900
tttgaatgga	atctaaatgc	aagtgaatct	ccagaaagca	acatagctac	gctgaccaca	960
tccccctttt	tcttttctta	cattacagaa	gacgacgaat	aa		1002

<210> 62

<211> 333

<212> PRT

-50-

<213> Artificial Sequence

<220>

<223> 35F RGD

<400> 62

```

Met Thr Lys Arg Val Arg Leu Ser Asp Ser Phe Asn Pro Val Tyr Pro
 1      5      10      15
Tyr Glu Asp Glu Ser Thr Ser Gln His Pro Phe Ile Asn Pro Gly Phe
 20      25      30
Ile Ser Pro Asn Gly Phe Thr Gln Ser Pro Asp Gly Val Leu Thr Leu
 35      40      45
Lys Cys Leu Thr Pro Leu Thr Thr Thr Gly Gly Ser Leu Gln Leu Lys
 50      55      60
Val Gly Gly Gly Leu Thr Val Asp Asp Thr Asp Gly Thr Leu Gln Glu
 65      70      75      80
Asn Ile Arg Ala Thr Ala Pro Ile Thr Lys Asn Asn His Ser Val Glu
 85      90      95
Leu Ser Ile Gly Asn Gly Leu Glu Thr Gln Asn Asn Lys Leu Cys Ala
100      105      110
Lys Leu Gly Asn Gly Leu Lys Phe Asn Asn Gly Asp Ile Cys Ile Lys
115      120      125
Asp Ser Ile Asn Thr Leu Trp Thr Gly Ile Asn Pro Pro Pro Asn Cys
130      135      140
Gln Ile Val Glu Asn Thr Asn Thr Asn Asp Gly Lys Leu Thr Leu Val
145      150      155      160
Leu Val Lys Asn Gly Gly Leu Val Asn Gly Tyr Val Ser Leu Val Gly
165      170      175
Val Ser Asp Thr Val Asn Gln Met Phe Thr Gln Lys Thr Ala Asn Ile
180      185      190
Gln Leu Arg Leu Tyr Phe Asp Ser Ser Gly Asn Leu Leu Thr Glu Glu
195      200      205
Ser Asp Leu Lys Ile Pro Leu Lys Asn Lys Ser Ser Thr Ala Thr Ser
210      215      220
Glu Thr Val Ala Ser Ser Lys Ala Phe Met Pro Ser Thr Thr Ala Tyr
225      230      235      240
Pro Phe Asn Thr Thr Thr Arg Asp Ser Glu Asn Tyr Ile His Gly Ile
245      250      255
Cys Tyr Tyr Met Thr Ser Tyr Asp Arg Ser Leu Phe Pro Leu Asn Ile
260      265      270
Ser Ile Met Leu Asn Ser Arg Met Ile Ser Ser Asn Val His Cys Asp
275      280      285
Cys Arg Gly Asp Cys Phe Cys Ala Tyr Ala Ile Gln Phe Glu Trp Asn
290      295      300
Leu Asn Ala Ser Glu Ser Pro Glu Ser Asn Ile Ala Thr Leu Thr Thr
305      310      315      320
Ser Pro Phe Phe Phe Ser Tyr Ile Thr Glu Asp Asp Glu
325      330

```

<210> 63

<211> 1164

<212> DNA

<213> Artificial Sequence

<220>

<223> 41sF

<400> 63

```

atgaaaagaa ccagaattga agacgacttc aaccccgctct acccctatga caccttctca 60
actcccagca tcccctatgt agctccgccc ttcgtttctt ctgacgggtt acaggaaaaa 120

```

-51-

```

ccccaggag ttttagcact caagtacact gaccccat  ctaccaatgc taagcatgag 180
cttactttta aacttggaag caacataact ttagaaaatg gggtactttc ggccacagtt 240
cccactgttt ctctcccct tacaacagct aacaactccc tgggttttagc cacatccgct 300
cccatagctg tatcagctaa ctctctcaca ttggccaccg ccgcaccact gacagtaagc 360
aacaaccagc ttagtattaa cgcggggcaga gggttagtta taactaacia tgccttaaca 420
gttaatccta ccggagcgct aggtttcaat aacacaggag ctttacaatt aaatgctgca 480
ggaggaatga gagtggacgg tgccaactta attcttcatg tagcatatcc ctttgaagca 540
atcaaccagc taacactgcg attagaaaac gggttagaag taaccagcgg aggaaagctt 600
aacgttaagt tgggatcagg cctccaattt gacagtaacg gacgcattgc tattagtaat 660
agcaaccgaa ctggaagtgt accatccctc actaccattt ggtctatctc gcctacgcct 720
aactgctcca tttatgaaac ccaagatgca aacctatttc tttgtctaac taaaaacgga 780
gctcacgtat taggtactat aacaatcaaa ggtcttaaag gagcactgcg ggaaatgcac 840
gataacgctc tatctttaaa acttcccttt gacaatcagg gaaatttact taactgtgcc 900
ttggaatcat ccacctggcg ttaccaggaa accaacgcag tggcctctaa tgccttaaca 960
tttatgccca acagtacagt gtatccacga aacaaaaccg ctcaccgggg caacatgctc 1020
atccaaatct cgcctaacat caccttcagt gtcgtctaca acgagataaa cagtgggtat 1080
gcttttactt ttaaattggtc agccgaaccg ggaaaacctt ttcaccacc taccgctgta 1140
ttttgctaca taactgaaga ataa
1164

```

<210> 64

<211> 387

<212> PRT

<213> Artificial Sequence

<220>

<223> 41sF

<400> 64

```

Met Lys Arg Thr Arg Ile Glu Asp Asp Phe Asn Pro Val Tyr Pro Tyr
1      5      10      15
Asp Thr Phe Ser Thr Pro Ser Ile Pro Tyr Val Ala Pro Pro Phe Val
20      25      30
Ser Ser Asp Gly Leu Gln Glu Lys Pro Pro Gly Val Leu Ala Leu Lys
35      40      45
Tyr Thr Asp Pro Ile Thr Thr Asn Ala Lys His Glu Leu Thr Leu Lys
50      55      60
Leu Gly Ser Asn Ile Thr Leu Glu Asn Gly Leu Leu Ser Ala Thr Val
65      70      75      80
Pro Thr Val Ser Pro Pro Leu Thr Asn Ser Asn Asn Ser Leu Gly Leu
85      90      95
Ala Thr Ser Ala Pro Ile Ala Val Ser Ala Asn Ser Leu Thr Leu Ala
100     105     110
Thr Ala Ala Pro Leu Thr Val Ser Asn Asn Gln Leu Ser Ile Asn Ala
115     120     125
Gly Arg Gly Leu Val Ile Thr Asn Asn Ala Leu Thr Val Asn Pro Thr
130     135     140
Gly Ala Leu Gly Phe Asn Asn Thr Gly Ala Leu Gln Leu Asn Ala Ala
145     150     155     160
Gly Gly Met Arg Val Asp Gly Ala Asn Leu Ile Leu His Val Ala Tyr
165     170     175
Pro Phe Glu Ala Ile Asn Gln Leu Thr Leu Arg Leu Glu Asn Gly Leu
180     185     190
Glu Val Thr Ser Gly Gly Lys Leu Asn Val Lys Leu Gly Ser Gly Leu
195     200     205
Gln Phe Asp Ser Asn Gly Arg Ile Ala Ile Ser Asn Ser Asn Arg Thr
210     215     220
Arg Ser Val Pro Ser Leu Thr Thr Ile Trp Ser Ile Ser Pro Thr Pro
225     230     235     240
Asn Cys Ser Ile Tyr Glu Thr Gln Asp Ala Asn Leu Phe Leu Cys Leu
245     250     255
Thr Lys Asn Gly Ala His Val Leu Gly Thr Ile Thr Ile Lys Gly Leu

```

-52-

Lys	Gly	Ala	Leu	Arg	Glu	Met	His	Asp	Asn	Ala	Leu	Ser	Leu	Lys	Leu
		275					280					285			
Pro	Phe	Asp	Asn	Gln	Gly	Asn	Leu	Leu	Asn	Cys	Ala	Leu	Glu	Ser	Ser
	290					295					300				
Thr	Trp	Arg	Tyr	Gln	Glu	Thr	Asn	Ala	Val	Ala	Ser	Asn	Ala	Leu	Thr
305					310					315					320
Phe	Met	Pro	Asn	Ser	Thr	Val	Tyr	Pro	Arg	Asn	Lys	Thr	Ala	His	Pro
			325					330						335	
Gly	Asn	Met	Leu	Ile	Gln	Ile	Ser	Pro	Asn	Ile	Thr	Phe	Ser	Val	Val
			340					345					350		
Tyr	Asn	Glu	Ile	Asn	Ser	Gly	Tyr	Ala	Phe	Thr	Phe	Lys	Trp	Ser	Ala
		355					360					365			
Glu	Pro	Gly	Lys	Pro	Phe	His	Pro	Pro	Thr	Ala	Val	Phe	Cys	Tyr	Ile
	370					375					380				
Thr	Glu	Glu													
385															

<210> 65
 <211> 1194
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> 41sF RGD

<400> 65
 atgaaaagaa ccagaattga agacgacttc aaccccgctct acccctatga caccttctca 60
 actcccagca tcccctatgt agctccgccc ttcgtttctt ctgacggggt acaggaaaaa 120
 cccccaggag ttttagcact caagtacact gaccccata ctaccaatgc taagcatgag 180
 cttactttta aacttggaag caacataact ttagaaaatg gggttacttt ggccacagtt 240
 cccactgttt ctccctccct tacaacacag aacaactccc tgggttttag ccatccgct 300
 cccatagctg tatcagctaa ctctctcaca ttggccaccg ccgcaccact gacagtaagc 360
 aacaaccagc ttagtattaa cgcgggcaga ggtttagtta taactaaca tgccttaaca 420
 gttaatccta ccggagcgct aggtttcaat aacacaggag ctttacaatt aaatgctgca 480
 ggaggaatga gagtggacgg tgccaactta attcttcag tagcatatcc ctttgaagca 540
 atcaaccagc taacactgcg attagaaaac gggtagaag taaccagcgg aggaaagctt 600
 aacgttaagt tgggatcagg cctccaattt gacagtaacg gacgcattgc tattagtaat 660
 agcaaccgaa ctcgaaagt accatccctc actaccattt ggtctatctc gcctacgcct 720
 aactgctcca tttatgaaac ccaagatgca aacctatttc tttgtctaac taaaaacgga 780
 gctcacgtat taggtactat aacaatcaaa ggtcttaaag gagcactgcg ggaaatgcac 840
 gataacgctc tatcttttaa acttcccttt gacaatcagg gaaatttact taactgtgcc 900
 ttggaatcat ccacctggcg ttaccaggaa accaacgcag tggcctctaa tgccttaaca 960
 tttatgcccc acagtacagt gtatccacga aacaaaaccg ctcacccggg caacatgctc 1020
 atccaaatct cgcctaacat caccttcagt gtcgtctaca acgagataaa ctgtgattgt 1080
 cgtgggtgatt gtttttgtac tagtgggtat gcttttactt ttaaattggtc agccgaaccg 1140
 ggaaaacctt ttcaccacc taccgctgta ttttgctaca taactgaaga ataa 1194

<210> 66
 <211> 397
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> 41sF RGD

<400> 66
 Met Lys Arg Thr Arg Ile Glu Asp Asp Phe Asn Pro Val Tyr Pro Tyr
 1 5 10 15
 Asp Thr Phe Ser Thr Pro Ser Ile Pro Tyr Val Ala Pro Pro Phe Val

-53-

Ser	Ser	Asp	Gly	Leu	Gln	Glu	Lys	Pro	Pro	Gly	Val	Leu	Ala	Leu	Lys
		35					40					45			
Tyr	Thr	Asp	Pro	Ile	Thr	Thr	Asn	Ala	Lys	His	Glu	Leu	Thr	Leu	Lys
	50					55					60				
Leu	Gly	Ser	Asn	Ile	Thr	Leu	Glu	Asn	Gly	Leu	Leu	Ser	Ala	Thr	Val
65					70					75					80
Pro	Thr	Val	Ser	Pro	Pro	Leu	Thr	Asn	Ser	Asn	Asn	Ser	Leu	Gly	Leu
				85					90					95	
Ala	Thr	Ser	Ala	Pro	Ile	Ala	Val	Ser	Ala	Asn	Ser	Leu	Thr	Leu	Ala
			100					105					110		
Thr	Ala	Ala	Pro	Leu	Thr	Val	Ser	Asn	Asn	Gln	Leu	Ser	Ile	Asn	Ala
		115					120					125			
Gly	Arg	Gly	Leu	Val	Ile	Thr	Asn	Asn	Ala	Leu	Thr	Val	Asn	Pro	Thr
	130					135					140				
Gly	Ala	Leu	Gly	Phe	Asn	Asn	Thr	Gly	Ala	Leu	Gln	Leu	Asn	Ala	Ala
145					150					155					160
Gly	Gly	Met	Arg	Val	Asp	Gly	Ala	Asn	Leu	Ile	Leu	His	Val	Ala	Tyr
				165					170					175	
Pro	Phe	Glu	Ala	Ile	Asn	Gln	Leu	Thr	Leu	Arg	Leu	Glu	Asn	Gly	Leu
			180					185					190		
Glu	Val	Thr	Ser	Gly	Gly	Lys	Leu	Asn	Val	Lys	Leu	Gly	Ser	Gly	Leu
		195					200					205			
Gln	Phe	Asp	Ser	Asn	Gly	Arg	Ile	Ala	Ile	Ser	Asn	Ser	Asn	Arg	Thr
	210				215						220				
Arg	Ser	Val	Pro	Ser	Leu	Thr	Thr	Ile	Trp	Ser	Ile	Ser	Pro	Thr	Pro
225					230					235					240
Asn	Cys	Ser	Ile	Tyr	Glu	Thr	Gln	Asp	Ala	Asn	Leu	Phe	Leu	Cys	Leu
				245					250					255	
Thr	Lys	Asn	Gly	Ala	His	Val	Leu	Gly	Thr	Ile	Thr	Ile	Lys	Gly	Leu
			260					265					270		
Lys	Gly	Ala	Leu	Arg	Glu	Met	His	Asp	Asn	Ala	Leu	Ser	Leu	Lys	Leu
		275					280					285			
Pro	Phe	Asp	Asn	Gln	Gly	Asn	Leu	Leu	Asn	Cys	Ala	Leu	Glu	Ser	Ser
		290				295					300				
Thr	Trp	Arg	Tyr	Gln	Glu	Thr	Asn	Ala	Val	Ala	Ser	Asn	Ala	Leu	Thr
305					310					315					320
Phe	Met	Pro	Asn	Ser	Thr	Val	Tyr	Pro	Arg	Asn	Lys	Thr	Ala	His	Pro
				325					330					335	
Gly	Asn	Met	Leu	Ile	Gln	Ile	Ser	Pro	Asn	Ile	Thr	Phe	Ser	Val	Val
			340					345					350		
Tyr	Asn	Glu	Ile	Asn	Cys	Asp	Cys	Arg	Gly	Asp	Cys	Phe	Cys	Thr	Ser
		355					360					365			
Gly	Tyr	Ala	Phe	Thr	Phe	Lys	Trp	Ser	Ala	Glu	Pro	Gly	Lys	Pro	Phe
	370					375					380				
His	Pro	Pro	Thr	Ala	Val	Phe	Cys	Tyr	Ile	Thr	Glu	Glu			
385					390					395					

<210> 67

<211> 1737

<212> DNA

<213> Artificial Sequence

<220>

<223> Ad5 PD1 penton

<400> 67

atgcggcgcg cggcgatgta tgaggaaggt cctcctcct cctacgagag tgtgggtgagc 60
 gggcgccag tggcgggcgc gctgggttct cccttcgatg ctcccctgga cccgccgttt 120
 gtgcctccgc ggtacctgcg gcctaccggg gggagaaaca gcatccgtta ctctgagttg 180

-54-

```

gcacccctat tcgacaccac ccgtgtgtac ctggtggaca acaagtcaac ggatgtggca 240
tccctgaact accagaacga ccacagcaac tttctgacca cggtcattca aaacaatgac 300
tacagcccgg gggaggcaag cacacagacc atcaatcttg acgaccgggc gactgggggc 360
ggcgacctga aaaccatcct gcataccaac atgccaaatg tgaacgagtt catgtttacc 420
aataagttta aggcgcgggt gatggtgtcg cgcttgccca ctaaggacaa tcaggtggag 480
ctgaaatacg agtgggtgga gttcacgctg cccgagggca actactccga gaccatgacc 540
atagacctta tgaacaacgc gatcgtggag cactacttga aagtgggcag acagaacggg 600
gttctggaaa gcgacatcgg ggtaaaagttt gacacccgca acttcagact ggggtttgac 660
cccgtcactg gtcttgtcat gcctggggta tatacaaacg aagccttcca tccagacatc 720
atcttgctgc caggatgcgg ggtggacttc acccacagcc gcctgagcaa cttgttgggc 780
atccgcaagc ggcaaccctt ccaggagggc tttaggatca cctacgatga tctggagggt 840
ggtaacattc ccgactgtt ggatgtggac gcctaccagg cgagcttgaa agatgacacc 900
gaacagggcg ggggtggcg aggcggcagc aacagcagtg gcagcggcg ggaagagAAC 960
tccaacgcgg cagccgcggc aatgcagccg gtggaggaca tgaacgatag ccgcggtac 1020
ccctacgacg tgcccgaacta cgccgggcacc agcgccacac gggctgagga gaagcgcgt 1080
gaggccgaag cagcggccga agctgccgcc cccgctgcgc aacccgaggt cgagaagcct 1140
cagaagaaac cggatgatcaa acccctgaca gaggacagca agaaacgcag ttacaaccta 1200
ataagcaatg acagcacctt caccacgtac cgcagctggg accttgcata caactacggc 1260
gacctcaga ccggaatccg ctcatggacc ctgctttgca ctctgacgt aacctgcggc 1320
tcggagcagg tctactggtc gttgccagac atgatgcaag acccgtgac cttccgtcc 1380
acgcgccaga tcagcaactt tccggtggtg ggcgccgagc tgttgcccgt gactccaag 1440
agcttctaca acgaccagge cgtctactcc caactcatcc gccagtttac ctctctgacc 1500
cacgtgttca atcgctttcc cgagaaccag attttggcgc gcccgccagc cccaccatc 1560
accacgctca gtgaaaacgt tctgtctctc acagatcaag ggacgctacc gctgcgcaac 1620
agcatcggag gagtccagcg agtgaccatt actgacgcca gacgccgcac ctgccctac 1680
gtttacaagg ccctgggcat agtctcgccg cgcgtcctat cgagccgcac tttttga 1737

```

<210> 68

<211> 578

<212> PRT

<213> Artificial Sequence

<220>

<223> Ad5 PD1 penton

<400> 68

```

Met Arg Arg Ala Ala Met Tyr Glu Glu Gly Pro Pro Pro Ser Tyr Glu
1      5      10
Ser Val Val Ser Ala Ala Pro Val Ala Ala Ala Leu Gly Ser Pro Phe
20     25     30
Asp Ala Pro Leu Asp Pro Pro Phe Val Pro Pro Arg Tyr Leu Arg Pro
35     40     45
Thr Gly Gly Arg Asn Ser Ile Arg Tyr Ser Glu Leu Ala Pro Leu Phe
50     55     60
Asp Thr Thr Arg Val Tyr Leu Val Asp Asn Lys Ser Thr Asp Val Ala
65     70     75     80
Ser Leu Asn Tyr Gln Asn Asp His Ser Asn Phe Leu Thr Thr Val Ile
85     90     95
Gln Asn Asn Asp Tyr Ser Pro Gly Glu Ala Ser Thr Gln Thr Ile Asn
100    105    110
Leu Asp Asp Arg Ser His Trp Gly Asp Leu Lys Thr Ile Leu His
115    120    125
Thr Asn Met Pro Asn Val Asn Glu Phe Met Phe Thr Asn Lys Phe Lys
130    135    140
Ala Arg Val Met Val Ser Arg Leu Pro Thr Lys Asp Asn Gln Val Glu
145    150    155    160
Leu Lys Tyr Glu Trp Val Glu Phe Thr Leu Pro Glu Gly Asn Tyr Ser
165    170    175
Glu Thr Met Thr Ile Asp Leu Met Asn Asn Ala Ile Val Glu His Tyr
180    185    190
Leu Lys Val Gly Arg Gln Asn Gly Val Leu Glu Ser Asp Ile Gly Val

```

-55-

Lys	Phe	195	Asp	Thr	Arg	Asn	Phe	200	Arg	Leu	Gly	Phe	Asp	205	Pro	Val	Thr	Gly
	210						215						220					
Leu	Val	Met	Pro	Gly	Val	Tyr	Thr	Asn	Glu	Ala	Phe	His	Pro	Asp	Ile			
225					230					235					240			
Ile	Leu	Leu	Pro	Gly	Cys	Gly	Val	Asp	Phe	Thr	His	Ser	Arg	Leu	Ser			
				245					250					255				
Asn	Leu	Leu	Gly	Ile	Arg	Lys	Arg	Gln	Pro	Phe	Gln	Glu	Gly	Phe	Arg			
			260					265					270					
Ile	Thr	Tyr	Asp	Asp	Leu	Glu	Gly	Gly	Asn	Ile	Pro	Ala	Leu	Leu	Asp			
	275						280					285						
Val	Asp	Ala	Tyr	Gln	Ala	Ser	Leu	Lys	Asp	Asp	Thr	Glu	Gln	Gly	Gly			
	290					295					300							
Gly	Gly	Ala	Gly	Gly	Ser	Asn	Ser	Ser	Gly	Ser	Gly	Ala	Glu	Glu	Asn			
305					310					315					320			
Ser	Asn	Ala	Ala	Ala	Ala	Ala	Met	Gln	Pro	Val	Glu	Asp	Met	Asn	Asp			
				325					330					335				
Ser	Arg	Gly	Tyr	Pro	Tyr	Asp	Val	Pro	Asp	Tyr	Ala	Gly	Thr	Ser	Ala			
			340					345					350					
Thr	Arg	Ala	Glu	Glu	Lys	Arg	Ala	Glu	Ala	Glu	Ala	Ala	Ala	Glu	Ala			
	355						360					365						
Ala	Ala	Pro	Ala	Ala	Gln	Pro	Glu	Val	Glu	Lys	Pro	Gln	Lys	Lys	Pro			
	370					375					380							
Val	Ile	Lys	Pro	Leu	Thr	Glu	Asp	Ser	Lys	Lys	Arg	Ser	Tyr	Asn	Leu			
385					390					395				400				
Ile	Ser	Asn	Asp	Ser	Thr	Phe	Thr	Gln	Tyr	Arg	Ser	Trp	Tyr	Leu	Ala			
				405					410					415				
Tyr	Asn	Tyr	Gly	Asp	Pro	Gln	Thr	Gly	Ile	Arg	Ser	Trp	Thr	Leu	Leu			
			420					425					430					
Cys	Thr	Pro	Asp	Val	Thr	Cys	Gly	Ser	Glu	Gln	Val	Tyr	Trp	Ser	Leu			
	435						440					445						
Pro	Asp	Met	Met	Gln	Asp	Pro	Val	Thr	Phe	Arg	Ser	Thr	Arg	Gln	Ile			
	450					455					460							
Ser	Asn	Phe	Pro	Val	Val	Gly	Ala	Glu	Leu	Leu	Pro	Val	His	Ser	Lys			
465					470					475				480				
Ser	Phe	Tyr	Asn	Asp	Gln	Ala	Val	Tyr	Ser	Gln	Leu	Ile	Arg	Gln	Phe			
			485					490						495				
Thr	Ser	Leu	Thr	His	Val	Phe	Asn	Arg	Phe	Pro	Glu	Asn	Gln	Ile	Leu			
	500						505						510					
Ala	Arg	Pro	Pro	Ala	Pro	Thr	Ile	Thr	Thr	Val	Ser	Glu	Asn	Val	Pro			
	515						520					525						
Ala	Leu	Thr	Asp	His	Gly	Thr	Leu	Pro	Leu	Arg	Asn	Ser	Ile	Gly	Gly			
	530					535					540							
Val	Gln	Arg	Val	Thr	Ile	Thr	Asp	Ala	Arg	Arg	Arg	Thr	Cys	Pro	Tyr			
545					550					555				560				
Val	Tyr	Lys	Ala	Leu	Gly	Ile	Val	Ser	Pro	Arg	Val	Leu	Ser	Ser	Arg			
				565					570					575				
Thr	Phe																	

<210> 69

<211> 1773

<212> DNA

<213> Artificial Sequence

<220>

<223> 5TS35H

<400> 69

atgaagcgcg caagaccgctc tgaagatacc ttcaaccccg tgtatccata tgacacggaa 60

-56-

```

accggtccctc caactgtgcc ttttcttact cctccctttg tatcccccaa tgggtttcaa 120
gagagtcctcc ctgggggtact ctctttgctc ctatccgaac ctctagttac ctccaatggc 180
atgcttgcgc tcaaaatggg caacggcctc tctctggagc aggcgggcaa ccttacctcc 240
caaaatgtaa ccactgtgag cccacctctc aaaaaaacca agtcaaacat aaacctggaa 300
atatctgcac ccctcacagt tacctcagaa gccctaactg tggctgcccgc cgcacctcta 360
atggctgcgg gcaacacact caccatgcaa tcacaggccc cgctaaccgt gcacgactcc 420
aaacttagca ttgccaccca aggaccctc acagtgtcag aaggaaagct agccctgcaa 480
acatcaggcc ccctcaccac caccgatagc agtaccctta ctactactgc ctcacccctc 540
ctaactactg ccactggtag cttgggcatt gacttgaaag agcccattta tacacaaaat 600
ggaaaactag gactaaagta cggggctcct ttgcatgtaa cagacgacct aaacactttg 660
accgtagcaa ctgggtccagg tgtgactatt aataatactt ccttgcaaac taaagtact 720
ggagccttgg gttttgattc acaaggcaat atgcaactta atgtagcagg aggactaagg 780
attgattctc aaaacagacg ccttatactt gatgttagtt atccgtttga tgctcaaac 840
caactaaatc taagactagg acagggccct ctttttataa actcagccca caacttggat 900
attaactaca acaaaggcct ttacttgttt acagcttcaa acaattccaa aaagcttgag 960
gttaacctaa gcactgcca ggggttgatg tttgacgcta cagccatagc cattaatgca 1020
ggagatgggc ttgaatttgg ttcacctaat gcaccaaaca caaatccctt caaaacaaaa 1080
attggccatg gcctagaatt tgattcaaac aaggctatgg ttcctaaact aggaactggc 1140
cttagttttg acagcacagg tgccattaca gtaggaaaca aaaataatga taagctaact 1200
ttgtggaccg gaataaaccc tccacctaac tgtcaaattg tggaaaacac taatacaaat 1260
gatggcaaac ttacttttagt attagtaaaa aatggagggc ttgttaatgg ctacgtgtct 1320
ctagttgggtg tatcagacac tgtgaaccaa atgttcacac aaaagacagc aaacatccaa 1380
ttaagattat attttgactc ttctggaaat ctattaactg aggaatcaga cttaaaaaat 1440
ccacttaaaa ataaatcttc tacagcgacc agtgaaactg tagccagcag caaagccttt 1500
atgccaagta ctacagctta tcccttcaac accactacta gggatagtga aaactacatt 1560
catggaatat gttactacat gactagttat gatagaagtc tatttcctct gaacatttct 1620
ataatgctaa acagccgtat gatttcttcc aatgttgctt atgccatata atttgaatgg 1680
aatctaaatg caagtgaatc tccagaaagc aacatagcta cgctgaccac atcccccttt 1740
ttcttttctt acattacaga agacgacgaa taa 1773

```

<210> 70

<211> 590

<212> PRT

<213> Artificial Sequence

<220>

<223> 5TS35H

<400> 70

```

Met Lys Arg Ala Arg Pro Ser Glu Asp Thr Phe Asn Pro Val Tyr Pro
1      5      10      15
Tyr Asp Thr Glu Thr Gly Pro Pro Thr Val Pro Phe Leu Thr Pro Pro
20      25      30
Phe Val Ser Pro Asn Gly Phe Gln Glu Ser Pro Pro Gly Val Leu Ser
35      40      45
Leu Arg Leu Ser Glu Pro Leu Val Thr Ser Asn Gly Met Leu Ala Leu
50      55      60
Lys Met Gly Asn Gly Leu Ser Leu Asp Glu Ala Gly Asn Leu Thr Ser
65      70      75      80
Gln Asn Val Thr Thr Val Ser Pro Pro Leu Lys Lys Thr Lys Ser Asn
85      90      95
Ile Asn Leu Glu Ile Ser Ala Pro Leu Thr Val Thr Ser Glu Ala Leu
100     105     110
Thr Val Ala Ala Ala Ala Pro Leu Met Val Ala Gly Asn Thr Leu Thr
115     120     125
Met Gln Ser Gln Ala Pro Leu Thr Val His Asp Ser Lys Leu Ser Ile
130     135     140
Ala Thr Gln Gly Pro Leu Thr Val Ser Glu Gly Lys Leu Ala Leu Gln
145     150     155     160
Thr Ser Gly Pro Leu Thr Thr Thr Asp Ser Ser Thr Leu Thr Ile Thr
165     170     175

```


-57-

Ala	Ser	Pro	Pro	Leu	Thr	Thr	Ala	Thr	Gly	Ser	Leu	Gly	Ile	Asp	Leu
			180					185					190		
Lys	Glu	Pro	Ile	Tyr	Thr	Gln	Asn	Gly	Lys	Leu	Gly	Leu	Lys	Tyr	Gly
		195					200					205			
Ala	Pro	Leu	His	Val	Thr	Asp	Asp	Leu	Asn	Thr	Leu	Thr	Val	Ala	Thr
	210					215					220				
Gly	Pro	Gly	Val	Thr	Ile	Asn	Asn	Thr	Ser	Leu	Gln	Thr	Lys	Val	Thr
	225				230					235					240
Gly	Ala	Leu	Gly	Phe	Asp	Ser	Gln	Gly	Asn	Met	Gln	Leu	Asn	Val	Ala
			245						250					255	
Gly	Gly	Leu	Arg	Ile	Asp	Ser	Gln	Asn	Arg	Arg	Leu	Ile	Leu	Asp	Val
			260					265					270		
Ser	Tyr	Pro	Phe	Asp	Ala	Gln	Asn	Gln	Leu	Asn	Leu	Arg	Leu	Gly	Gln
		275					280					285			
Gly	Pro	Leu	Phe	Ile	Asn	Ser	Ala	His	Asn	Leu	Asp	Ile	Asn	Tyr	Asn
	290				295						300				
Lys	Gly	Leu	Tyr	Leu	Phe	Thr	Ala	Ser	Asn	Asn	Ser	Lys	Lys	Leu	Glu
	305				310					315					320
Val	Asn	Leu	Ser	Thr	Ala	Lys	Gly	Leu	Met	Phe	Asp	Ala	Thr	Ala	Ile
				325					330					335	
Ala	Ile	Asn	Ala	Gly	Asp	Gly	Leu	Glu	Phe	Gly	Ser	Pro	Asn	Ala	Pro
			340					345					350		
Asn	Thr	Asn	Pro	Leu	Lys	Thr	Lys	Ile	Gly	His	Gly	Leu	Glu	Phe	Asp
		355					360					365			
Ser	Asn	Lys	Ala	Met	Val	Pro	Lys	Leu	Gly	Thr	Gly	Leu	Ser	Phe	Asp
	370					375					380				
Ser	Thr	Gly	Ala	Ile	Thr	Val	Gly	Asn	Lys	Asn	Asn	Asp	Lys	Leu	Thr
	385				390					395					400
Leu	Trp	Thr	Gly	Ile	Asn	Pro	Pro	Pro	Asn	Cys	Gln	Ile	Val	Glu	Asn
				405					410					415	
Thr	Asn	Thr	Asn	Asp	Gly	Lys	Leu	Thr	Leu	Val	Leu	Val	Lys	Asn	Gly
			420					425					430		
Gly	Leu	Val	Asn	Gly	Tyr	Val	Ser	Leu	Val	Gly	Val	Ser	Asp	Thr	Val
		435					440					445			
Asn	Gln	Met	Phe	Thr	Gln	Lys	Thr	Ala	Asn	Ile	Gln	Leu	Arg	Leu	Tyr
	450					455					460				
Phe	Asp	Ser	Ser	Gly	Asn	Leu	Leu	Thr	Glu	Glu	Ser	Asp	Leu	Lys	Ile
	465				470					475					480
Pro	Leu	Lys	Asn	Lys	Ser	Ser	Thr	Ala	Thr	Ser	Glu	Thr	Val	Ala	Ser
			485						490					495	
Ser	Lys	Ala	Phe	Met	Pro	Ser	Thr	Thr	Ala	Tyr	Pro	Phe	Asn	Thr	Thr
			500					505					510		
Thr	Arg	Asp	Ser	Glu	Asn	Tyr	Ile	His	Gly	Ile	Cys	Tyr	Tyr	Met	Thr
		515					520					525			
Ser	Tyr	Asp	Arg	Ser	Leu	Phe	Pro	Leu	Asn	Ile	Ser	Ile	Met	Leu	Asn
	530					535					540				
Ser	Arg	Met	Ile	Ser	Ser	Asn	Val	Ala	Tyr	Ala	Ile	Gln	Phe	Glu	Trp
	545				550					555					560
Asn	Leu	Asn	Ala	Ser	Glu	Ser	Pro	Glu	Ser	Asn	Ile	Ala	Thr	Leu	Thr
			565						570					575	
Thr	Ser	Pro	Phe	Phe	Phe	Ser	Tyr	Ile	Thr	Glu	Asp	Asp	Glu		
			580					585					590		

<210> 71

<211> 945

<212> DNA

<213> Artificial Sequence

<220>

<223> 35TS5H

-58-

<400> 71

```

atgaccaaga gagtccggct cagtgactcc ttcaaccctg tctaccacct tgaagatgaa 60
agcacctccc aacaccocct tataaaaccca gggttttattt ccccaaattg cttcacacaa 120
agcccagacg gagttctttac tttaaaatgt ttaaccccac taacaaccac aggcggatct 180
ctacagctaa aagtgggagg gggacttaca gtggatgaca ctgatggtac cttacaagaa 240
aacatacgtg ctacagcacc cattactaaa aataatcact ctgtagaact atccattgga 300
aatggattag aaactcaaaa caataaacta tgtgccaaat tgggaaatgg gttaaaattt 360
aacaacggtg acatttgtat aaaggatagt attaacacct tatggactac accagctcca 420
tctcctaact gtagactaaa tgcagagaaa gatgctaaac tcacttttgg ctttaacaaaa 480
tgtggcagtc aaatacttgc tacagtttca gttttggctg tttaaaggcag tttggctcca 540
atatctggaa cagttcaaa tgctcatctt attataagat ttgacgaaaa tggagtgcta 600
ctaaacaatt ccttcctgga cccagaatat tggaaacttta gaaatggaga tcttactgaa 660
ggcacagcct atacaaacgc tgttggattt atgcctaacc tatcagctta tccaaaatct 720
cacggtaaaa ctgccaaaag taacattgtc agtcaagttt acttaaacgg agacaaaaat 780
aaacctgtaa cactaacctt tacactaaac ggtacacagg aaacaggaga cacaactcca 840
agtgcatact ctatgtcatt ttcattgggac tgggtctggcc acaactacat taatgaaata 900
tttgccacat cctcttacac tttttcatac attgcccag aataa 945

```

<210> 72

<211> 314

<212> PRT

<213> Artificial Sequence

<220>

<223> 35TS5H

<400> 72

```

Met Thr Lys Arg Val Arg Leu Ser Asp Ser Phe Asn Pro Val Tyr Pro
 1      5      10      15
Tyr Glu Asp Glu Ser Thr Ser Gln His Pro Phe Ile Asn Pro Gly Phe
 20     25     30
Ile Ser Pro Asn Gly Phe Thr Gln Ser Pro Asp Gly Val Leu Thr Leu
 35     40     45
Lys Cys Leu Thr Pro Leu Thr Thr Thr Gly Gly Ser Leu Gln Leu Lys
 50     55     60
Val Gly Gly Gly Leu Thr Val Asp Asp Thr Asp Gly Thr Leu Gln Glu
 65     70     75     80
Asn Ile Arg Ala Thr Ala Pro Ile Thr Lys Asn Asn His Ser Val Glu
 85     90     95
Leu Ser Ile Gly Asn Gly Leu Glu Thr Gln Asn Asn Lys Leu Cys Ala
100    105    110
Lys Leu Gly Asn Gly Leu Lys Phe Asn Asn Gly Asp Ile Cys Ile Lys
115    120    125
Asp Ser Ile Asn Thr Leu Trp Thr Thr Pro Ala Pro Ser Pro Asn Cys
130    135    140
Arg Leu Asn Ala Glu Lys Asp Ala Lys Leu Thr Leu Val Leu Thr Lys
145    150    155    160
Cys Gly Ser Gln Ile Leu Ala Thr Val Ser Val Leu Ala Val Lys Gly
165    170    175
Ser Leu Ala Pro Ile Ser Gly Thr Val Gln Ser Ala His Leu Ile Ile
180    185    190
Arg Phe Asp Glu Asn Gly Val Leu Leu Asn Asn Ser Phe Leu Asp Pro
195    200    205
Glu Tyr Trp Asn Phe Arg Asn Gly Asp Leu Thr Glu Gly Thr Ala Tyr
210    215    220
Thr Asn Ala Val Gly Phe Met Pro Asn Leu Ser Ala Tyr Pro Lys Ser
225    230    235    240
His Gly Lys Thr Ala Lys Ser Asn Ile Val Ser Gln Val Tyr Leu Asn
245    250    255
Gly Asp Lys Thr Lys Pro Val Thr Leu Thr Ile Thr Leu Asn Gly Thr
260    265    270

```

-59-

Gln	Glu	Thr	Gly	Asp	Thr	Thr	Pro	Ser	Ala	Tyr	Ser	Met	Ser	Phe	Ser
		275					280					285			
Trp	Asp	Trp	Ser	Gly	His	Asn	Tyr	Ile	Asn	Glu	Ile	Phe	Ala	Thr	Ser
	290					295					300				
Ser	Tyr	Thr	Phe	Ser	Tyr	Ile	Ala	Gln	Glu						
305					310										